University of Alberta

Developments Towards Manufacturable Nanoelectronics and Applications

by

Jarosław Mikołaj Sulima ©

Submitted to the Faculty of Graduate Studies and Research in partial fulfillment

of the requirements for the degree of Master of Science

Department of Electrical and Computer Engineering

Edmonton, Alberta

Fall 2004

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-95861-2
Our file Notre référence
ISBN: 0-612-95861-2

The author has granted a non-exclusive license allowing the Library and Archives Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ACKNOWLEDGEMENTS

First and foremost I wish to acknowledge the unconditional support of my parents and my sister. Without their support, both financial and spiritual, this thesis would not have been possible.

I would also like to thank my supervisor, Dr. Chris Backhouse, for taking me on and for his ever upbeat and positive outlook. Thanks also go out to iCore and Dr. Michael Brett for their roles in funding this research.

Many thanks are due to Brian Moore, who not only allowed me to use his equipment, but spent several nights with me testing and experimenting with it.

Finally, I would like to thank Holly Rourke, for being such a good friend and to José, Ranjit, Eric and the entire Backhouse lab for creating a positive work environment.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1 CMOS AND NANOELECTRONICS 1

1.2 OVERVIEW OF THESIS 5

CHAPTER 2: ELECTRON TUNNELLING THEORY 7

2.1 CHAPTER OVERVIEW 7

2.2 CLASSICAL MODEL 7

2.3 CONDUCTION REGIMES 8

2.4 TUNNELLING CURRENT 11

2.5 THEORETICAL MODELS 14

2.6 EFFECT OF OXIDE ROUGHNESS 18

CHAPTER 3: PREVIOUS WORK AND APPLICATIONS 19

3.1 CHAPTER OVERVIEW 19

3.2 POINT CONTACT MIM DIODES 19

3.3 THIN-FILM MIM DIODES 20

3.4 APPLICATIONS OF TUNNELLING DEVICES 25

3.4.1 Metal-Insulator Tunnelling Transistor (MITT) *26*

3.4.2 Mixing *29*

3.4.3 Harmonic Radar *30*

3.4.4 Rectification *32*

3.4.5 Terahertz Imaging *35*

3.4.6 Magnetic Tunnel Junctions *37*

3.4.7 Superconducting Tunnel Junctions *41*

CHAPTER 4: FABRICATION OF TUNNEL JUNCTIONS **48**

4.1 CHAPTER OVERVIEW 48

4.2 MIM JUNCTIONS 48

4.3 LCD 55

4.4 SIS JUNCTIONS 57

CHAPTER 5: MIM DIODE PROCESS DEVELOPMENT **61**

5.1 CHAPTER OVERVIEW 61

5.2 MIM DESIGN 61

5.3 CROSS-LAYER MIM DIODE PROCESS 63

5.4 PHYSICAL CHARACTERIZATION 69

5.4.1 Section Overview *69*

5.4.2 Metal Characterization *70*

5.4.3 Photolithography *71*

5.4.4 Ellipsometry *76*

CHAPTER 6: D.C RESULTS AND ANALYSIS **89**

6.1 CHAPTER OVERVIEW 89

6.2 DC TESTING OF MIM DIODES 89

6.3 YIELD 95

6.3.1 Run 1 .... *98*

6.3.2 Run 2 .... *98*

6.3.3 Run 3 .... *99*

6.3.4 Run 4 .... *102*

6.3.5 Run 5 .... *102*

6.4 FITS OF THEORY WITH MEASURED DATA 104

6.5 SUMMARY 107

CHAPTER 7: TUNNELLING TRANSISTOR DESIGN 109

7.1 CHAPTER OVERVIEW 109

7.2 TRANSISTOR DESIGN 109

7.3 PROCESS STEPS 117

7.4 MITT MATERIALS AND GEOMETRY CONSIDERATIONS 121

CHAPTER 8: R.F. CHARACTERIZATION AND HARMONIC RADAR TAG DESIGN 126

8.1 CHAPTER OVERVIEW 126

8.2 R.F. MEASUREMENT OF DIODE PROPERTIES 126

8.3 TAG DESIGN 132

8.4 ANTENNA DESIGN 134

CHAPTER 9: CONCLUSIONS AND FUTURE WORK 144

9.1 CHAPTER OVERVIEW 144

9.2 MAJOR CHALLENGES 145

9.3 FUTURE WORK 146

9.4 SUMMARY 150

BIBLIOGRAPHY 151

LIST OF TABLES

Table 1 Resistances of measured diodes 94

Table 2 Summary of all sputtering run parameters 97

Table 3 Summary of e-beam run parameters (deposition rates are for top electrode) 97

Table 4 Summary of devices in highest yield wafer 101

Table 5 Angled evaporation effect of variables on tunnel barrier geometry 116

Table 6 Conversion efficiency vs. input power for MIM devices at RF frequencies 129

Table 7 Substrate effects on antenna parameters. Only electrically thick (relative to wavelength) substrates have effect on resonance wavelength and impedance does not vary much. 141

LIST OF FIGURES

Figure 1 Current regimes in a metal-insulator-metal structure 9

Figure 2 An F-N plot of the mid-voltage region of Simmons' model (presented in Section 2.5). The higher voltage (lower 1/V) approaches a straight line and the F-N tunnelling regime. The high voltage model itself yields a simple negatively sloped straight line. 11

Figure 3 Band diagrams of the conduction regimes of a tunnel junction a) no applied voltage b) intermediate voltage c) high voltage 15

Figure 4 Rounding effect of the image force on the insulating barrier 17

Figure 5 Cross layer metal-oxide-metal junction 22

Figure 6 Edge metal-oxide-metal junction 23

Figure 7 Approximate *i-v* curve of a Ni-NiO-Ni diode built by Havemman.[26] 25

Figure 8 Two variations of an MITT transistor (Matsumura group[40,41]) 27

Figure 9 Theoretical basis of operation of the MITT is insulator conduction band deformation by the applied gate voltage 27

Figure 10 An MITT fabricated with STM oxidation schematic (from Snow *et al.* [39]) 28

Figure 11 Mixing of 10.6um lasers by an MIM diode (Masalmeh *et al.*[27]) 30

Figure 12 A simple harmonic radar device: folded dipole antenna (top) 31

Figure 13 Rectenna used for the detection of laser signals (Hoofring *et al.*[8]) 33

Figure 14 An MRAM 'cross-point' architecture with CMOS (*Tehrani et al*), top ferromagnetic layer is free, bottom is fixed. 40

Figure 15 Typical *i-v* characteristics of a Josephson Junction (underdamped) with I_0, the critical current of the superconductors 42

Figure 16 Schematic of a D.C. SQUID 44
Figure 17 Schematic of an R.F.-SQUID 45
Figure 18 A lift-off process a) photoresist is spun on a wafer b) the surface of the photoresist is depleted of photoactive compounds by soaking in chlorobenzene c) the sample is patterned with a photomask and UV light d) an overhang results on the sidewall d) metal is deposited. 52
Figure 19 MIM-LCD of Chrome, Tantalum Oxide and Tantalum (left) and addressing schematic of an MIM pixel (right) 55
Figure 20 The Nb/Al/AlOx/Nb structure created by Veidt *et al*[78] 60
Figure 21 Four point probing of a thin sample 70
Figure 22 Image of an MIM diode with a patterned 5μm (2μm mask pattern) second electrode layer of photoresist as seen through an optical microscope (left) and chronic pattern underexposure for a 5μm (1μm mask pattern) line (right). 73
Figure 23 Overhead view of a 3μm undercut. Outer pattern is LOR, inner pattern is photoresist on top of the LOR 75
Figure 24 Isometric view of overhanging photoresist 75
Figure 25 Optical image of completed MIM diode 76
Figure 26 Elliptically polarized light from reflection off a sample 77
Figure 27 Diagram of a typical laser ellipsometer 79
Figure 28 Reflection of light at thin-film on substrate 79
Figure 29 Psi-Delta curve for NiO on Ni from 0 to 1850Å, as calculated from equations 33,38-41. 83
Figure 30 Psi-Delta curve (zoomed) to 0 to 53Å (each cross = 1Å) 83

Figure 31 NiO removal from sputter deposited Ni by Ar^+ plasma (107W) exposure in an RIE chamber 86

Figure 32 Pressure and power dependence of the $Ar^+:O_2$ sputter-etch oxidation technique 88

Figure 33 D.C. Characterization set-up 90

Figure 34 Tunnel *i-v* diode characteristics from run 1 (~30Å oxide thickness) before and after dielectric breakdown........ 91

Figure 35 *i-v* characteristics of 40Å (nominal) barrier diodes from run 3 92

Figure 36 Fowler-Nordheim plots of selected diodes. The diodes asymptotically approach the F-N regime (linear, negative slope F-N plot characteristics)........ 93

Figure 37 Diodes from run 5 (with the LOR resist layer unremoved). A malformed device (left) has had its top Ni layer removed. A diode with well-defined features (right)104

Figure 38 Selected fits of data with the theory of Simmons JA= junction area in m^2, phi = work function difference between metal and barrier in eV, d = barrier length in Å (clockwise from top left diodes are 7M, 7L, 7E and 6M) 106

Figure 39 A curve fit using Simmons' theory by Scherer *et al*[31] of a Cr/CrOx/Cr tunnel junction 107

Figure 40 Process steps for an MITT fabricated with anodic oxidation in ethyline glycol-dissolved ammonium tetraborate a) Formation of gate insulator b) Ti evaporation and lithography c) Dry etching of half part of Ti d) Anodic oxidation of Ti e) Evaporation of Ti f) Lift-off process[40] 110

Figure 41 MITT process flow for a deposited tunnel barrier........ 112

Figure 42 Geometry of the overlapping source-drain tunnel junction method........ 112

Figure 43 Slightly overlapping electrodes by angled evaporation (left) and plasma oxidation creation of tunnel junction (right) 114

Figure 44 Evaporation angle non-uniformity (ΔΦ) 115

Figure 45 Finite-element simulation of the electrostatic characteristics of an ideal straight-walled MITT, showing deformed field in the tunnel region with applied gate bias. a) cross-section of MITT, b) gate voltage = 0V, c) gate voltage = 2.5V , d) gate voltage = 5V, e) gate voltage = 0V and relative dielectric permittivity of gate lowered to 4. 124

Figure 46 Measurement of MIM diode RF characteristics. 128

Figure 47 Spectrum scan of harmonic generation by the MIM diode at 9-11MHz excitation and 3dBm applied power. 131

Figure 48 Harmonic generation at 100MHz excitation and 9dBm applied power 132

Figure 49 Unpopulated (PCB) R.F. tag by Moore/Backhouse 134

Figure 50 λ/2 folded dipole design parameters 136

Figure 51 Input impedance (top-real, bottom-imaginary) of a folded dipole antenna.... 138

Figure 52 Asymmetric coplanar strip line (with a finite dielectric thickness) geometry 140

Appendix:

Figure A - 1 MIMSLF Mask 167

Figure A - 2 MIMSBOE Mask 168

Figure A - 3 MIMSSF Mask 169

Figure A - 4 Folded-Dipole Harmonic Radar Tag Antenna: The antenna is specified to the dimensions discussed in Chapter 8. The diode is a simple cross-layer configuration, similar to the MIM diode process. 170

LIST OF ABBREVIATIONS AND SYMBOLS

α-Si	Amorphous Silicon
AES	Auger Electrons Spectroscopy
AMR	Anisotropic Magnetoresistance
BOE	Buffered Oxide Etch
CMOS	Complementary Metal Oxide Semiconductor
DOS	Density of States
DI	De-ionized
EDX	Electron Dispersive X-Ray (Analysis)
ESD	Electrostatic Discharge
FM	Ferromagnet, Ferromagnetic
FN	Fowler-Nordheim
GMR	Giant Magnetoresistance
JA	Junction Area
J_{fn}	Fowler-Nordheim Current Density
J_{th}	Current Density At Temperatures > 0K
k	Optical Extinction Coefficient
IPA	Isopropyl Alcohol
KOH	Potassium Hydroxide
LCD	Liquid Crystal Display
LOR	Lift-Off Resist
MEMS	Microelectromechanical Systems
MIM	Metal-Insulator-Metal (Junction)

MITT	Metal-Insulator Tunnelling Transistor
MOSFET	Metal-Oxide Field Effect Transistor
n	Optical Refractive Index
ψ	Electronic Quantum Wavefunction
RBS	Rutherford Backscattering (Analysis)
RIE	Reactive Ion Etching
RF	Radio Frequency
RFID	Radio Frequency Identification
RMS	Root Mean Squared
RSFQ	Rapid Single Flux Quantum
SDT	Spin-Dependent Tunnelling
SEM	Scanning Electron Micrograph
Si	Silicon
SiO_2	Silicon Dioxide
SIMS	Secondary Ion Mass Spectroscopy
SIS	Superconductor-Insulator-Superconductor (Junction)
SQUID	Superconducting Quantum Interference Device
STM	Scanning Tunnelling Microscopy/Microscope
TEM	Transmission Electron Microscope
TMAH	Tetra Methyl Ammonium Hydroxide
TMR	Tunnelling Magnetoresistance
UV	Ultraviolet
θ	Tunnelling Probability

1

Chapter 1: Introduction

1.1 CMOS and Nanoelectronics

Since the late 1980s the MOSFET has been the dominant switching device in integrated electronics. In the search for higher operational frequency and integration density the MOSFET has been scaled down to nanometre dimensions. The semiconductor industry is planning to implement minimum gate length dimensions of just 25nm by 2007.[1] At the time of this writing, gate oxide thicknesses are on the order of a single nanometre. At these small dimensions quantum effects play a significant role in the operation of the MOSFET and, in general, they are undesired effects that limit performance. There exists a theoretical limit on the functional usability of the MOSFET that lies somewhere in the single digit nanometre gate length range (with off-state leakage current due to quantum tunnelling as the cause), but even before this limit is reached gate-drain tunnelling through the gate insulator is expected to challenge scaling.[2,3] Apart from the quantum effects, there are the considerable challenges of lithography, uniformity, higher electric fields, the short channel effect and vanishing bulk properties of the nanometre size MOSFETs as well as the rapidly increasing costs of fabricating smaller devices that would have to be addressed.[4,5] Because of these limitations it is reasonable to speculate that the next era of integrated electronics will feature devices that will rely on quantum mechanical effects for operation rather than regard them as 'noise'.

Quantum devices have several advantages over the MOSFET, in addition to their superior nanometre scalability. Since many quantum devices are fabricated using only deposited metals and insulators the cost to fabricate the devices is minimal. The cost of current state of the art CMOS manufacturing infrastructure increases by 60%[5] (per wafer) every 3 years requiring new facilities that cost billions of dollars. At this rate, there is sufficient motivation to research alternate devices (to CMOS) purely due to economic reasons.

Currently, MEMS and CMOS electronics are fabricated separately, requiring mutually exclusive fabrication facilities. The two chips are then connected in a process known as flip-chipping. This is a somewhat awkward solution, since different fabrication facilities are necessary to manufacture each chip. Nanoelectronics, made with standard MEMS process, could solve this integration problem by fabricating both device types on a single wafer. Also, since most quantum devices can be made with surface micromachining techniques they are potentially much simpler to fabricate than MOSFETs and this may allow them to be integrated into three-dimensional circuits.

Finally, quantum devices rely on shorter transition distances and generally have smaller capacitances, therefore they may also allow higher frequencies of operation (into

THz frequencies). These reasons, and the fact that the body of work on quantum devices pales in comparison to that of the MOSFET, motivate further research.

There are several quantum devices currently being researched, such as: quantum dots (QD), the single electron transistor (SET), resonant tunnelling transistor (RTT) and diode (RTD), Josephson junctions and the metal-insulator-metal (MIM) diode. They all share the same principle method of operation: electron tunnelling. With the only other notable nanoelectronic device in research being in the chemical science domain (i.e. molecular electronics) these tunnelling devices represent the key research areas in nanoelectronics. Since there are not yet practical methods of interfacing with molecular electronics it is clear that any research towards *practical* nanoelectronics should focus on microfabricated tunnelling devices. Considering this, some of the tunnelling devices, such as QDs and the SET, require minimum feature sizes that are too small for current practical fabrication methods (10nm or less for room temperature operation). The MIM diode, however, *is* manufacturable using conventional photolithographic techniques.

While a justified goal in its own right, nanoelectronics is not the only possible use of tunnel junctions. The MIM diode has a wealth of possible applications, both current and future, to which research can be applied. This diode has been historically researched as a high-frequency mixer and harmonic generator. It exhibits non-linear characteristics up to very high frequencies (even up to optical frequencies, depending on junction area capacitance) and have led to many experiments into ultra-high frequency mixing, detection and frequency generation. In the form of superconducting Josephson junctions (superconductor-insulator-superconductor or SIS junctions) the device is currently the most sensitive low-noise mixer of THz radiation, and has consequently found application in radioastronomy. Even non-superconducting metal tunnel junctions have been shown to be sensitive up to optical frequencies at room temperature. This holds promise for fields such as THz imaging, radiofrequency (R.F.) tracking and even infrared (I.R.) and solar rectification. Recently, magnetic forms of the MIM diode have been receiving much attention for application in non-volatile random access memory.[6] The magnetic tunnel junctions show the highest resistance change to magnetic field (magnetoresistance) of any technology and have only a few technical impediments to commercially viable manufacture. Clearly, the MIM diode is a versatile device.

The focus of this research is the design, fabrication and characterization of MIM diodes so that future research can be expanded to more complex tunnel junction-based nanoelectronic devices or the diodes can be integrated into practical manufactured systems. The fabrication of these tunnel junctions has enabled the design of two applications: metal tunnel junction transistors and R.F. harmonic radar tags that would eventually enable tracking and environmental monitoring with millimetre size devices. The proposed process for the tunnel junction transistor is for a metal insulator tunnel transistor (MITT), which relies on electrostatic field and tunnel effects. The harmonic radar tag design represents the first stage of integration of an R.F. tag system with the ultimate goal being a "smart-dust" sensor (i.e. mm scale) that can be interrogated remotely.

1.2 Overview of Thesis

This thesis is divided into nine chapters, including this chapter, the introduction. A discussion of the theory of quantum tunnelling is presented in Chapter 2. The governing equations of tunnel junctions are introduced, and, where applicable, derived from quantum theory.

Chapter 3 discusses previous work on the microfabricated MIM diode and devices based on MIM junctions. Several applications are discussed, such as use of the MIM diode for high frequency mixing, in an R.F. or microwave system as components of a harmonic radar tag, for detection (rectification) of electromagnetic waves up to near infrared frequencies, in magnetic memory and as superconducting junctions.

Chapter 4 discusses further details of the fabrication processes of prior work in tunnel junctions. Specifically, techniques, characteristics and challenges encountered in the fabrication of conventional metal MIM diodes, liquid crystal display (LCD) MIM diodes and SIS diodes are presented.

Chapter 5 describes the method of fabrication of the MIM diodes of the current research including the physical characterization of the devices and of the fabrication process.

Chapter 6 presents the results of the electrical characterization of the tunnelling devices and the interpretation of the data through the theory presented in Chapter 2. The diodes were tested at D.C. and the data points were compared with a tunnel junction model.

Chapter 7 discusses the design and microfabrication of a three terminal tunnelling device, the MITT transistor, using processes that are similar to those used in the construction of the MIM diode.

Chapter 8 presents the measured R.F. harmonic generation characteristics of the MIM diode developed here. This chapter also presents the design of a harmonic radar tag that incorporates the MIM diode. A simple antenna-diode system is designed that can be scaled to smaller dimensions and higher frequencies for ultimate application in a 'smart-dust' type tag.

Chapter 9 summarizes the research and includes a discussion of possible extension of the research in future work.

2

Chapter 2: Electron Tunnelling Theory

2.1 Chapter Overview

This chapter presents a theoretical background to the function of tunnelling devices, specifically MIM junctions. The derivation of a current-voltage model is followed, and the Simmons model for symmetric tunnel junctions is presented.

2.2 Classical Model

In classical electrodynamics, conduction occurs when a charge carrier, electron or hole, moves through a conducting or semi-conducting medium until it encounters a non-conducting barrier or volume of similar charge, at which point the electron stops or is reflected. This simple model excludes the possibility of conduction through an insulating medium unless an electron has sufficient energy to surmount the energy barrier of the insulator. In atomic terms, the electron must have enough energy to occupy a conduction band energy state in the insulator while it passes through it. At a temperature of 0K, no electrons possess this energy and, while metals can still conduct at this temperature, semiconductors and insulators cannot. As the temperature rises some electrons are thermally excited into higher energy levels and some electrons may have enough energy to occupy states in the conduction band of the insulator so that some conduction over the insulator barrier takes place. This type of conduction is known as thermal or thermionic current. In the non-quantum model, all of the current between two metals separated by an insulating layer is due to thermionic current. This is not the case in reality, as there can exist a large current in a metal-insulator-metal (MIM) junction at cryogenic temperatures. The phenomenon of current in an MIM junction (and other conductor-insulator-conductor junctions where the insulator is very thin) is explained by electron tunnelling. An electron, incident on a potential barrier, has a finite probability of penetrating a certain distance into the barrier. If the barrier thickness is small then it is possible for the electron to penetrate through the barrier entirely. A thinner barrier is associated with a higher probability of the electron tunnelling through the barrier in this manner. The electron can therefore surmount the barrier, through thermionic emission, or penetrate through the barrier by electron tunnelling and the total current across a tunnel junction, such as an MIM junction, is the sum of the tunnel current and the thermionic current.

2.3 Conduction Regimes

The tunnelling current through insulating barriers can be sub-divided into tunnelling due to direct tunnelling and that due to Fowler-Nordheim tunnelling. Though the tunnelling mechanism is the same for both regimes, the energies and electrical barrier widths are different. The distinction is best shown in Figure 1.

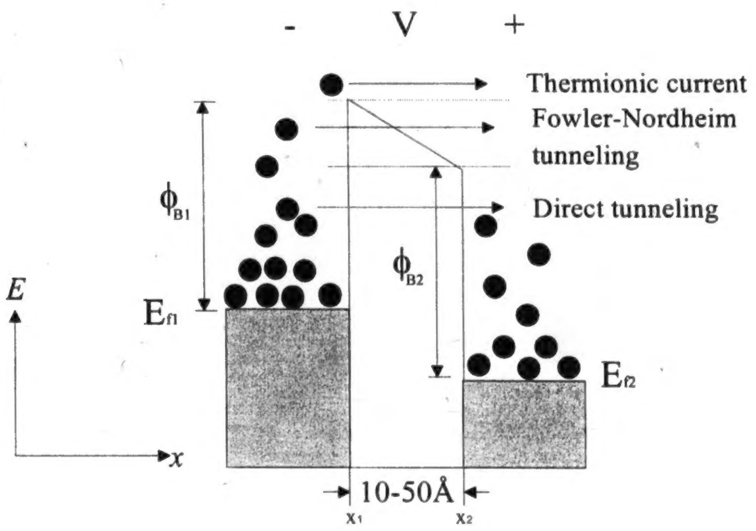


Figure 1 Current regimes in a metal-insulator-metal structure

The thermionic current is composed of electrons that have enough energy to surmount the energy barrier. It is generally regarded as an undesired component of the total current since it has a weak relationship with the bias voltage and it is an inherently slow process compared to tunnelling.[7,8] The direct tunnelling current is composed of electron travel between energy states on either electrode that are below the conduction band of the insulator. The current in this regime is approximately linear with applied voltage. As a result, this current regime is of secondary interest compared to the Fowler-Nordheim regime. The Fowler-Nordheim region is composed of electrons that tunnel through the sloped region of the insulator barrier, and hence through an effective insulator barrier width that is variable with applied voltage. As will be shown in the next section, the probability that an electron can tunnel through a barrier has an exponential dependence on the barrier width. The current in this regime is therefore very strongly dependent on the applied voltage and the current-voltage characteristics become very non-linear at high bias (V>ϕ_B, where ϕ_B is the barrier height at the metal/insulator interface) and the current density increases by orders of magnitude over the low bias case. Tunnelling current density in the Fowler-Nordheim regime is given by:[9]

$$J = \frac{2.2e^3E^2}{8\pi h\phi_B}\exp\left(\frac{-8\pi}{2.96heE}\sqrt{2m\phi_B^3}\right) \tag{1}$$

where e is the electronic charge, E is the electric field strength, h is Planck's constant, m is the electronic mass and ϕ_B is the barrier height in eV.

Most tunnelling devices rely on Fowler-Nordheim current to dominate over thermionic and direct tunnelling current. The insulator barrier thickness must be kept very small (below approximately 5nm) for this to be true for applied electric fields that do not induce dielectric breakdown.

Often it is difficult to prove that electron tunnelling is a dominant conduction mechanism in a device. One method to verify that the device is a tunnelling device is to graphically check the exponential relationship between the current and voltage. This is normally done in a Fowler-Nordheim (F-N) plot. A F-N plot is a plot of $\ln\left(\frac{I}{E^2}\right) \text{vs} \frac{1}{E}$. As can be quickly checked against the F-N current equation (above), this plot should yield a straight line with a negative slope. In this manner, if an F-N plot of data is a straight line then it is confirmed that the current equation has the form $J = AE^2\exp\left(-\frac{B}{E}\sqrt{\phi_B^3}\right)$, where A, B and ϕ_B are constants. At mid-voltage ranges the plot increases at a decreasing rate as 1/V increases (Figure 2). It is generally accepted that to confirm tunnelling as the dominant current, the data plot should be consistent with this shape of plot. Further, the F-N plot can be used to extract the value of the work-function ϕ_B if the effective tunnel thickness and effective junction area are known, *a-priori*.

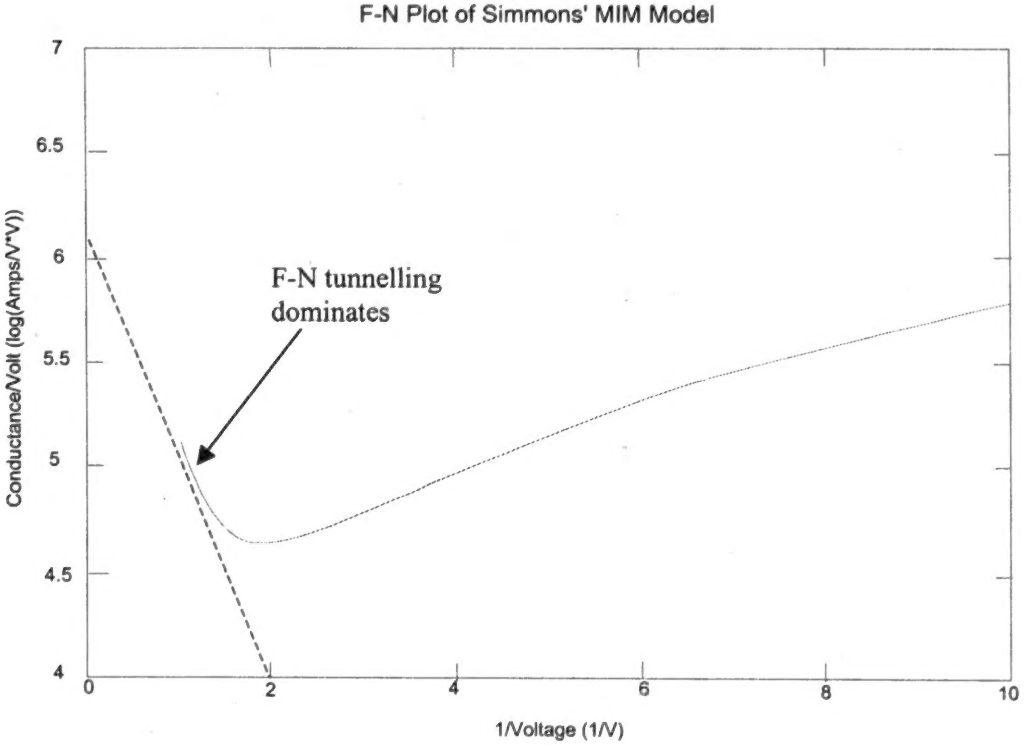


Figure 2 An F-N plot of the mid-voltage region of Simmons' model (presented in Section 2.5). The higher voltage (lower 1/V) approaches a straight line and the F-N tunnelling regime. The high voltage model itself yields a simple negatively sloped straight line.

2.4 Tunnelling Current

To determine the tunnelling probability of an electron it is necessary to describe the probability of finding an electron at a certain position. Due to the wave-particle duality of electrons a free electron's energy state $\Psi(\vec{r},t)$ can be described in terms of position and time by the wave equation:[10]

$$\Psi(\vec{r},t) = Ae^{i(\vec{k}\cdot\vec{r}-\omega t)} \tag{2}$$

where $\vec{r}$ represents displacement, $\vec{k}$ is the electron wave vector (related to the electron's momentum by $k=2\pi p/h$, h is Planck's constant, p is the electron momentum) and ω is the angular frequency of the electron wave. But to find the state of an electron in an arbitrary potential $U(x)$ the Schrödinger equation must be solved. The time-independent Schrödinger equation is given by:[11]

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x)}{\partial x}+U(x)\Psi(x)=E\Psi(x) \qquad (3)$$

The solution to this relationship is complex for systems where $U(x)$ is not constant so an approximation must be made. For a slowly varying potential change relative to the spatial rate of change of the wavelength λ ($dU(x)/dx \ll d\lambda/dx$) the wave equation solution to the Schrödinger equation is given by the Wentzel, Kramers, Brillouin (WKB) approximation:[12]

$$\Psi(L)=\Psi(0)e^{-\int_0^d k dx} \qquad (4)$$

where $k=\frac{\sqrt{2m[U(x)-E]}}{\hbar}$ and d is the thickness of the insulating barrier. The relative probability, $P(x)$, of finding an electron at a point in space is given by:[13]

$$P(x)=\Psi(x)\Psi^{*}(x) \qquad (5)$$

The tunnelling probability, $D(E)$, though a triangular barrier is just the probability that an electron exists at the far end of the barrier divided by the probability that it exists at the near end[17], using equation (4):

$$D(E)=\frac{\Psi(d)\Psi^{*}(d)}{\Psi(0)\Psi^{*}(0)}=\exp\left(-2\int_{0}^{d}\frac{\sqrt{2m(U(x)-E)}}{\hbar}dx\right) \tag{6}$$

$$D(E)\approx\exp\left(\frac{-4\pi d\sqrt{2m(U(x)-E)}}{h}\right) \tag{7}$$

$$D(E)\approx\exp\left(-4\pi d\frac{\sqrt{2m(E_{f}+\phi-E)}}{h}\right) \tag{8}$$

Where d is the insulator thickness, E_f is the Fermi energy and ϕ is the average barrier height across d. To determine the current density of thermionic current and tunnelling current at temperatures above 0K, it is necessary to determine the electron population as a function of energy. The probability of electrons occupying available states at a certain energy E is given by the Fermi-Dirac distribution function:[14]

$$f(E)=\frac{1}{1+e^{\frac{E-E_F}{k_B T}}} \tag{9}$$

where E_F is the Fermi energy, defined as the energy level at which the probability of an electron occupying an available state is ½, and k_B is Boltzmann's constant.

Assuming an ideal metal-insulator-metal junction (smooth surfaces, sharp transitions, image forces neglected etc.) with an insulator thickness on the order of a few tens of Angstroms and low temperatures (where thermionic current can be ignored) the tunnel current can be readily derived. The tunnel current consists of a summation of all electrons with a momentum $k_{\perp}$, the component of momentum in the direction perpendicular to the plane of the tunnel barrier, multiplied by the probability of tunnelling at that momentum. The tunnel current from electrode one to two is then:[15]

$$J_{12}=\frac{2mq}{h^{2}}\int\int D(k_{\perp})(f_{1}(E)-f_{2}(E))dk_{\perp}dE \tag{10}$$

Assuming a slowly varying potential $U(x)$ with respect to the spatial rate of change of the electron wavelength, the probability that an electron can tunnel through a potential barrier is given by: [16]

$$D(E_x) = \exp\left(\frac{-4\pi}{h} \int_{x1}^{x2} \left| \sqrt{2m(U(x) - E_x)} \right| dx \right) \tag{11}$$

where E_x is the energy component of incident electrons in the x direction and $U(x)$ is the height of the potential barrier at distance x from the edge of the barrier.

The number of electrons contributing to tunnelling across the insulating barrier is then just the number of electrons with a kinetic energy component E_x that is less than the height of the potential barrier. With a change of variables from momentum to energy and accounting for an applied voltage V in the electrodes, the equation becomes: [16]

$$J_{12} = \frac{4\pi mq}{h^3} \int_0^{E_m} D(E_x) \int_{E_x}^{\infty} \left(f(E) - f(E + eV) \right) dE \cdot dE_x \tag{12}$$

where dE is the total energy of the electrons. This is the tunnelling current density in one direction only. A similar term J_{21} accounts for the reverse current, though this current is considered negligible in most cases. From this starting point Simmons and others researchers have constructed their models of conduction.

2.5 Theoretical Models

The most common set of models used in research to model symmetric tunnel junctions is that proposed by Simmons. [9,17,18] Along with previous researchers Simmons divided the current-voltage relationships into three regions: the low, intermediate and high voltage ranges (Figure 3). The equations in this section can be found in "Generalized

Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film."[17]

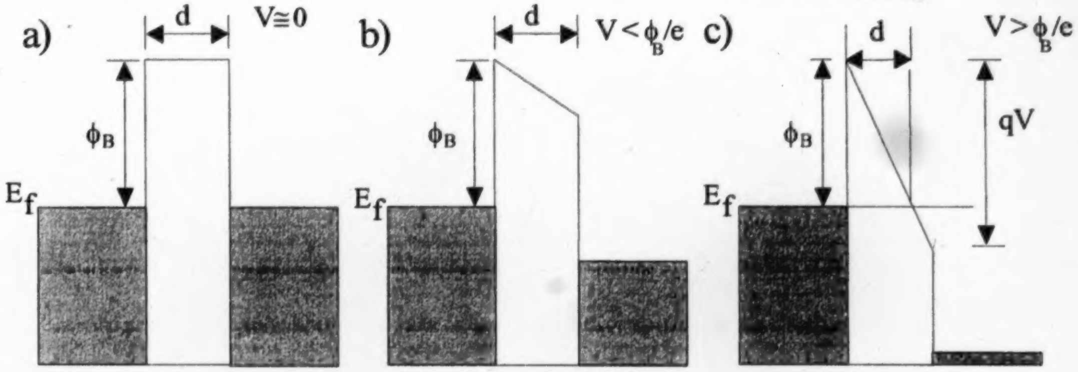


Figure 3 Band diagrams of the conduction regimes of a tunnel junction a) no applied voltage b) intermediate voltage c) high voltage

At low voltage, $V \approx 0$, the current density is a linear function of applied voltage:

$$J = \left(\frac{\sqrt{2m \cdot \phi_B}}{d}\right) \cdot \left(\frac{e}{h}\right)^2 \cdot V \cdot \exp\left(-\frac{4\pi \cdot \beta \cdot d \cdot \sqrt{2m}}{h}\right) \tag{13}$$

where d is the barrier thickness, V is the applied voltage, and β is a correction factor ≈ 1.

In the intermediate voltage range where $V < \phi_B / e$:

$$J = \left(\frac{e}{2\pi h \cdot d^2}\right) \cdot \left\{\left(\phi_B - \frac{eV}{2}\right)\exp\left[\frac{-4\pi d}{h}\sqrt{2m \cdot \left(\phi_B - \frac{eV}{2}\right)}\right] - \left(\phi_B + \frac{eV}{2}\right)\exp\left[\frac{-4\pi d}{h}\sqrt{2m \cdot \left(\phi_B + \frac{eV}{2}\right)}\right]\right\} \tag{14}$$

While in the high voltage range $V > \phi_B / e$ the current-voltage relationship takes the form of the Fowler-Nordheim equation (the first term represents the forward current while the second represents the reverse current which can be neglected for $V \gg \phi_B / e$):

$$J_{FN} = \frac{2.2e^2E^2}{8\pi h \cdot \phi_B}\left\{\exp\left(-\frac{8\pi}{2.96h \cdot e \cdot E}\sqrt{2m \cdot \phi_B^3}\right) - \left(1 + \frac{2e \cdot V}{\phi_B}\right)\exp\left(-\frac{8\pi}{2.96h \cdot e \cdot F}\sqrt{2m \cdot \phi_B^3\left(1 + \frac{2e \cdot V}{\phi_B}\right)}\right)\right\} \tag{15}$$

Simmons accounted for higher temperature (non-cryogenic) operation by including a thermal product term, defined by:

$$\frac{J(V,T)}{J(V,0)} = \frac{\pi BkT}{\sin \pi BkT} \simeq \left\{1 + \frac{1}{6}(\pi BkT)^2\right\} \tag{16}$$

However, this accounting for thermal distribution of electrons seems not to take into account the Schottky emission over the tunnel barrier (as it is a tunnelling model). The model for thermionic current is given by the Schottky current relation[19]:

$$J_{Schottky} = A_r \cdot T^2 \cdot \exp\left(\frac{-q\phi_B}{kT}\right) \cdot \left(\exp\left(\frac{qV_{Bias}}{kT}\right) - 1\right) \tag{17}$$

where A_r is the Richardson constant, given, for metals, by:

$$A_r = \frac{4\pi qmk^2}{h^3} \tag{18}$$

and V_{bias} is the voltage between the two electrodes through which Schottky current flows.

One final effect is accounted for when modelling tunnel junctions. The image force, due to the influence of electric charge on the potential barrier, is often included. The image force distorts the barrier from the sharp edged boundaries shown in Figure 3, to the rounded barrier boundary shown in Figure 4.

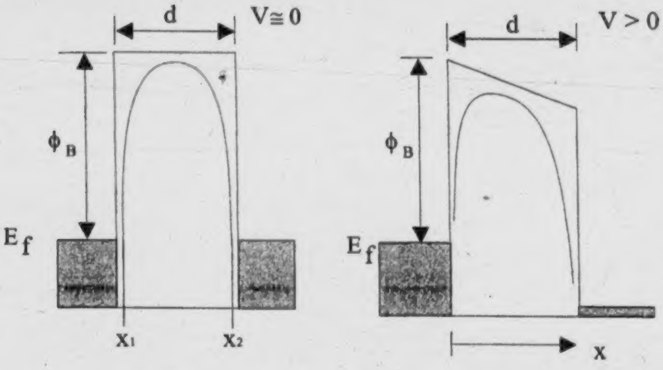


Figure 4 Rounding effect of the image force on the insulating barrier

Mathematically, Simmons found that the image force can be modeled accurately by:

$$\phi(x)=\phi_0 - eV\frac{x}{d} - \frac{1.15\lambda d^2}{x(d-x)} \tag{19}$$

which, when integrated across the width of the barrier yields an average barrier height, given by:

$$\phi_I=\phi_0-\frac{eV}{2d}(x_1+x_2)-\left(\frac{1.15\lambda d}{x_2-x_1}\right)\ln\left(\frac{x_2(d-x_1)}{x_1(d-x_2)}\right) \tag{20}$$

where x_1 and x_2 are the left and right limits of the insulating barrier (defined as the point on the conduction band of the insulator equal to the metal Fermi level). The average barrier height is then used in the current density equation:

$$J=J_0\left(\phi_I\exp\left(-A\cdot\phi_I^{1/2}\right)-\left(\phi_I+eV\right)\exp\left(-A\left(\phi_I+eV\right)^{1/2}\right)\right) \tag{21}$$

$$A=\frac{4\pi d}{h}\sqrt{2m} \tag{22}$$

$$J_0 = \frac{e}{2\pi h(\beta d)^2} \tag{23}$$

where β is a correction factor, less than, but close to 1.

This model by Simmons is the most frequently used model for symmetric tunnel junctions since it is the most complete and analytical solution for tunnel junction current. An alternate model, by Brinkman, Dynes and Rowell[20] (BDR) which also includes asymmetric junctions is also available,. However, the BDR model must be solved numerically and, as such, is more complex and awkward to deal with if only symmetric junctions are being modeled. For this reason, and due to its use in other works, the Simmons model is used in this research.

2.6 Effect of Oxide Roughness

The insulator thickness cannot be assumed to be the same as nominal due to statistical oxide thickness variation across the junction area. The statistical variation is a result of the roughness of the metal surface and inhomogeneous oxide growth. Due to the strong dependence of tunnel current on barrier thickness the tunnel current is dominated by the current at the thinnest regions of the insulator, known as hot-spots.[21] These regions are a small fraction of the physical junction area and the insulator thickness at these points can be several times smaller than the nominal or average thickness.[22] The measured values for oxide thickness can then be taken as a guide to device characteristics but cannot be used to directly model the device mathematically. The effective barrier thickness and effective tunnel junction area are often kept as free parameters when fitting to the data.

3

Chapter 3: Previous Work and Applications

3.1 Chapter Overview

A common and understandable question regarding this work may be ”why pursue research on the MIM tunnel junction diode?” This chapter presents the prior work on the MIM diode and several applications, which are either in the commercial (SIS junctions), research (magnetic junctions, tunnelling transistors, high frequency detectors and mixers) or concept stage (terahertz imaging is proposed here as a new application for the MIM diode). While the applications presented represent just a selection of possible uses for the MIM diode, each justifies research on the MIM diode in its own right.

The prior work on point-contact and thin-film diodes is first presented, followed by the applications of the MIM diode: the metal insulator tunnelling transistor (MITT), mixing, harmonic radar and rectification, THz imaging, magnetic junctions for magnetic field detection and non-volatile memory, and SIS junctions for magnetic field detection and low noise mixing.

3.2 Point Contact MIM Diodes

The metal-insulator-metal diode consists of two metal electrodes separated by a very thin insulator (on the order of a few nanometres). The insulating barrier is typically an oxide on one of the metal electrodes, as thin layers of insulating oxides grow naturally on most metals when exposed to oxygen. Due to its fast response characteristics the device has been in use since the 1970s for mixing and detection of sub-millimetre and infrared wavelength radiation in the form of a point-contact diode.[23,24] Point-contact diodes are made by gently touching a sharp metal tip to a polished metal layer covered by a natural oxide. The metal tip is most often an electrochemically etched metal whisker, typically tungsten etched in NaOH.[23,25] Because the physical basis for operation of the point contact diode is electron tunnelling and because the junction area is small (small capacitance), the device is responsive up to optical frequencies. It is a more sensitive detector than Schottky barrier diodes above about 12 THz[23], and can generate frequencies up to 200THz, at which point it has been theorized that the limitation stems not from the diode itself, but rather from poor electromagnetic coupling due to the skin-effect on the whisker antenna.[25,29]

While giving excellent response to high frequencies, the point-contact diode is generally unpredictable from contact to contact.[26] The AFM type diodes are inherently confined to the lab, and along with the whisker type point contact diodes, have been deemed too mechanically unstable for any practical non-academic use.[7,27,28] For this reason thin-film MIM diodes were developed.

3.3 Thin-Film MIM Diodes

While not typically functioning up to the high frequencies that the point-contact diodes are capable of, thin-film MIM diodes produced by various research groups have been mechanically stable. The reason for the poorer RF performance of the thin-film MIM diodes in comparison with point-contact diodes is that the contact areas, or junction areas, for the thin-film diodes are much larger. The larger areas correspond to a higher capacitance, and consequently, to a lower maximum frequency of response. The contact area for a typical MIM thin-film diode is about 0.5-10 μm^2, while a typical area for the point contact diode was estimated to be 0.02 μm.[2,7] Point-contact diodes can operate at over a hundred THz[29] while microfabricated tunnel junctions are generally outperformed by at least an order of magnitude in terms of frequency response.

Microfabricated MIM diodes can be categorized into two main types: the cross-layer MOM (metal-oxide-metal) diode and the edge MOM diode. The cross-layer MOM diode, shown in Figure 5, simply consists of a metal line that has been covered with an insulating layer and subsequently overlaid with another metal line. This configuration has been used by Wilke *et al.*[7] to create an MIM diode contact area of 0.056μm^2 and oxide thickness of 35Å. Their diode was used for 10.6μm wavelength CO_2 laser radiation detection. To achieve the small diode junction area the group used direct-write electron-beam lithography by patterning metal lines as small as 240nm in width. While electron beam lithography can pattern features as small as 10-50nm it is inherently a serial process (i.e. features are written one at a time). This is ideal for simple designs of a few devices, but for designs with many devices spread over a large area the time required to write the features becomes impractical. For manufacturability it is desirable to construct many devices on a wafer concurrently such that the cost per device is minimized, particularly if the ultimate device yield is poor. Electron beam nano-lithography is also typically confined to small areas (i.e. 100μm × 100μm for the University of Alberta Nanofab e-beam system) and it is a very expensive alternative for fabricating a wafer of an array of devices. It is therefore not viable for integration with MEMS processes.

The Edge-MOM diode, as developed by Heiblum *et al.*[28] and Havemann[26] can be made to have an extremely small contact area while being developed entirely with photolithographic tools. A representation of the edge-MOM diode is found in Figure 6. The contact area is defined by the vertical dimension of the bottom electrode and the width of the top electrode. This is ideal since material thickness is easily controlled to tens of nanometres precision with routine deposition, yielding a contact area as small as $0.01\mu m^2$. In comparison, the cross-layer MOM junction area made with the same photolithographic tools is at least $1\mu m^2$. With a bottom electrode thickness of only 100Å and a top electrode width of $1\mu m$ Heiblum's group achieved a $0.01\mu m^2$ contact area for their Ni-NiO-Ni diodes (though they published results only for $0.03\mu m^2$ contact areas and greater).

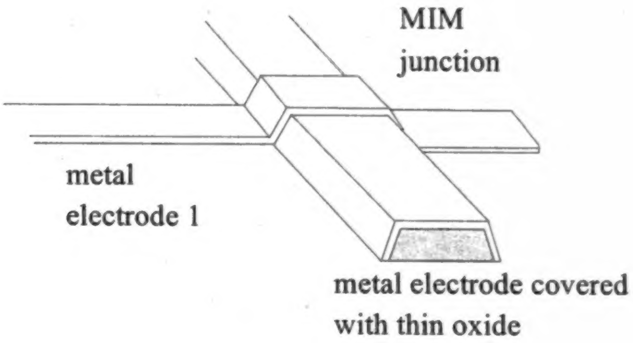


Figure 5 Cross layer metal-oxide-metal junction

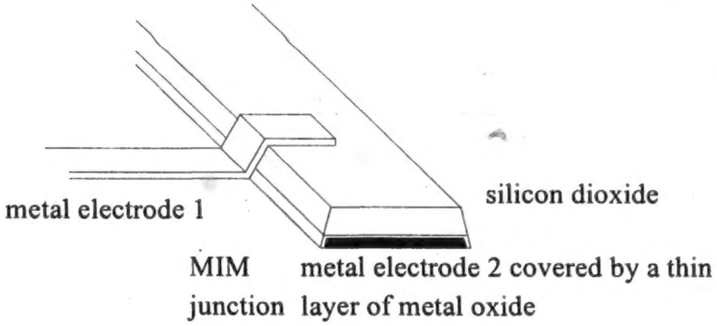


Figure 6 Edge metal-oxide-metal junction

The metal used by most groups researching conventional MIM diodes (i.e. diodes that were not fabricated specifically for magnetic or superconducting uses) was nickel as it was found to yield the best results.[26] It is not explicitly clear why nickel-nickel oxide-nickel diodes had better characteristics, but since diode behaviour is largely determined by oxide quality, it can be reasoned that nickel oxidation yields a relatively uniform and stable oxide layer. A range of other metals have also been used to construct tunnel junctions: aluminium[22,30] and chromium[31] for symmetric MIM diodes, and niobium and lead alloys, as well as Nb/Al/AlOx/Nb structures for SIS (Josephson) junction diodes. Asymmetric diodes, used for rectification and detection experiments, have been constructed with Ni/NiO/Au and Nb/NbO/Ag junctions.[8,32]

The quality of the oxide has been a critical issue in the development of tunnel junctions, particularly in SIS junctions where there has been the most research. Various investigations have shown that for high junction quality the barrier should be uniform, deposited on a smooth surface and constructed of a stable oxide.[33] Unfortunately, a high quality insulator is difficult to grow or deposit. Metallic shorts across the insulator (known as pin-holes) can exist in tunnelling barriers and decrease the quality of the desired junction characteristics (non-linearity, low sub-threshold current). Even if an insulating film is free from pinholes, the conduction could be dominated by current through hot-spots, small regions of very thin oxide, which could ultimately compromise the long-term stability of the junction since large currents are concentrated in small areas. Pinholes and hot-spots are generally the result of rough surfaces, therefore a uniform smooth surface is desired for deposition or growth of the insulator.

Generally, deposited oxides are not as uniform as grown oxides so most tunnel junctions are fabricated by growing an oxide on a metal electrode. Most recent SIS work, for example, presently concentrates on the Nb/Al-AlOx//Nb (e.g. Aoyagi *et al.*[34], Janawadkar *et al.*[36], Patel *et al.*[35]) structure which has a stable oxide that is uniform, and it has been reported that the Al wets the Nb electrode thereby not propagating the sputtered Nb grain structure and yielding a smoother surface.[36,78] For similar reasons AlOx is also commonly used in magnetic tunnel junctions (tunnel junctions that vary conductivity with magnetic orientation) as well. Unfortunately, the use of AlOx as the insulator barrier requires that the sample be kept under vacuum between the oxide growth and subsequent deposition, since Al oxidizes very rapidly in air to a thickness of 5nm. Various studies have been performed to determine insulator quality of tunnel junctions, but still, the quality of tunnel junctions remains largely a result of heuristic experimentation.

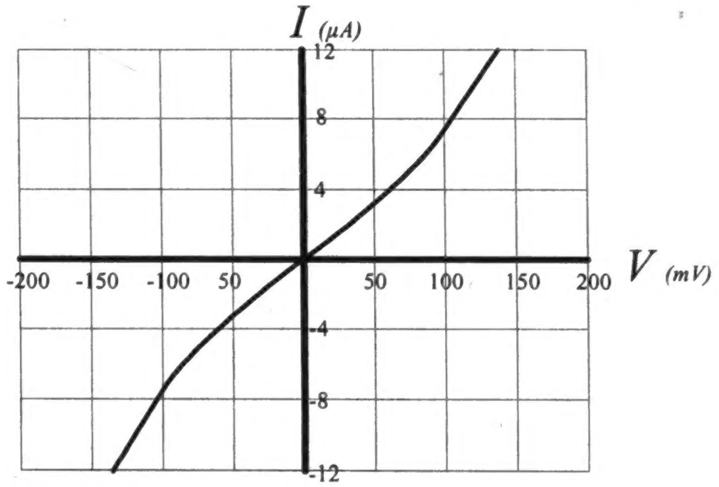


Figure 7 Approximate *i-v* curve of a Ni-NiO-Ni diode built by Havemman.[26]

While the *i-v* characteristics of diodes developed by different groups vary greatly, some trends do emerge. Most obviously, all the diodes have a low voltage linear region and a higher voltage non-linear region. A typical i-v characteristic for Havemann's edge diode is given in Figure 7. Although, judging from his 'typical' *i-v* curve, his diodes were not very non-linear; Havemann had successfully used his diodes to mix CO_2 lasers (28.3THz), to generate harmonics at MHz frequencies and to mix at GHz frequencies.

3.4 Applications of Tunnelling Devices

Practical applications of tunnelling devices already exist and future applications of tunnel junctions seem very promising. While the possible applications are numerous they can be grouped into two basic categories: tunnel junctions in electronic devices and tunnel junctions as high-frequency non-linear elements. The applications presented here are the metal-insulator tunnelling transistor (MITT), high frequency mixing, harmonic radar, high-frequency ac signal rectification, THz imaging, magnetic memory and in superconducting circuits.

3.4.1 Metal-Insulator Tunnelling Transistor (MITT)

The problem with most nanoelectronic transistor devices is that their processing methods are not compatible with manufacturable microfabrication technology. Tools such as electron beam lithography and focused ion beam lithography and even AFM probe tip techniques are necessary for the fabrication of small features of 10nm, but these tools are serial, "one feature at a time", in nature and not suitable for practical device fabrication. The fabrication of practical nanoelectronic devices (devices that can be fabricated readily in large numbers and that can be easily integrated with other components) currently requires the use of conventional photolithography with minimum feature sizes of micrometer scale (in most MEMS processes). It is desired to fabricate a practical nanoelectronic device that can be constructed with present photolithography capabilities but can also be scaled down to the nano-scale with future lithographic technology. A device that has been recently presented, the metal-insulator tunnelling transistor (MITT), is an ideal device for this development. The device is a natural extension of the MIM diode.

While the concept of a gate modulated tunnel barrier transistor most likely originated with a proposed Schottky barrier tunnelling diode by Tucker *et al*[37], the first MITT was proposed by Fujimaru and Matsumura[38] in 1996. It was subsequently verified in operation by Snow *et al.*[39] then by Matsumura's own group in 1998 and 1999 respectively.[40] The principle of operation of the device is control of tunnelling current by deformation of the shape of the tunnel barrier. This is done by applying an electric potential to a gate that is connected capacitively to a thick tunnel junction. The similarity to the MIM diode can be seen in Figure 8 (fabrication details in Chapter 8). The MITT is, in essence, an MIM diode with a gate oxide and electrode.

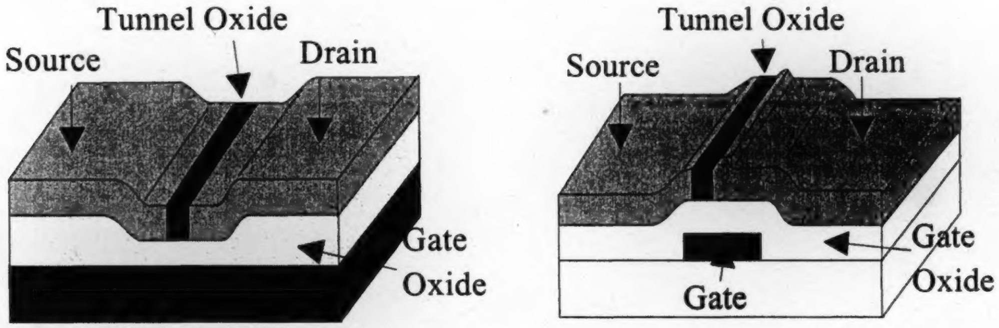

Tunnel Oxide
Source
Drain
Gate
Oxide
Tunnel Oxide
Source
Drain
Gate
Oxide
Gate

Figure 8 Two variations of an MITT transistor

The tunnel insulator width, W, is kept large (16 – 30nm) so as to suppress direct tunnelling current, which would be present in the off state as a leakage current. The gate oxide is thin (several nm), but has a high dielectric breakdown electric field and high work-function relative to both the gate and source/drain metals to minimize gate-source/drain tunnelling.

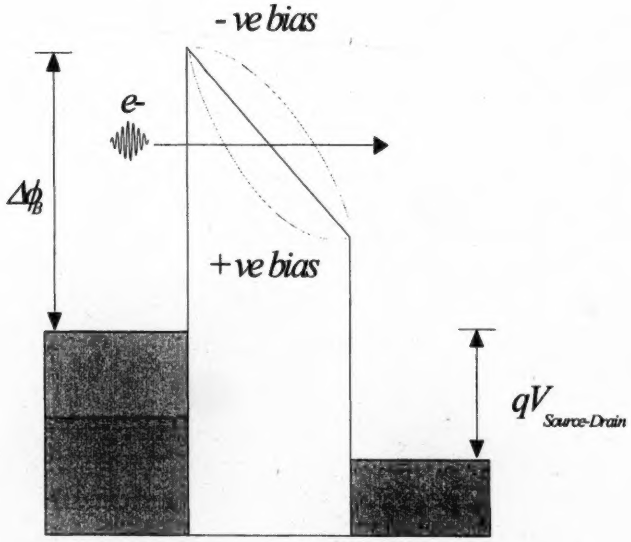

- ve bias
e-
$\Delta\phi_B$
+ ve bias
$qV_{Source\text{-}Drain}$

Figure 9 Theoretical basis of operation of the MITT is insulator conduction band deformation by the applied gate voltage

Figure 9 illustrates the theory of operation of the MITT. Since tunnelling current is very strongly a function of effective barrier thickness the modulation of the shape of the Fowler-Nordheim tunnelling barrier region by a gate voltage has a large effect on the total current across the source/drain junction.

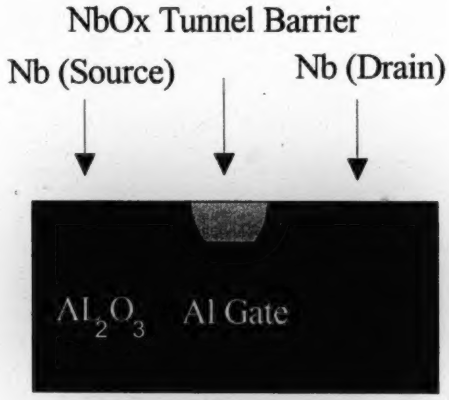


Figure 10 An MITT fabricated with STM oxidation schematic (from Snow *et al.* [39])

An MITT, fabricated by Snow *et al.*[39] (Figure 10), gave reasonable switching characteristics for a tunnelling device at 300K. The device was fabricated by oxidizing an aluminum gate using electrochemical anodic oxidation. A line of Nb was then deposited on top of the oxidized gate at a perpendicular angle and subsequently a tunnel oxide barrier was formed by scanning a conducting-tip AFM probe across the width of the Nb line following the path of the gate electrode below, at an applied voltage in an air atmosphere. MITT transistors have been made by at least three separate methods[39,40,41] with the most commercially viable of the three involving simple metal evaporation for all terminals and RIE etching, anodic oxidation and lift-off for the tunnel junction (Fujimaru *et al.*[40]). However, no fabrication process has yet been presented that produced devices using a completely manufacturable process.

3.4.2 Mixing

The MIM diode itself has the unique characteristic of having the same non-linear *i-v* characteristics from D.C. to as high as the infrared and optical spectrum (ultimately limited by its junction capacitance and at ultra high frequencies by the electron relaxation time).[109] This makes it an ideal detector of high frequency radiation through mixing. The MIM diode has been previously used to detect a 10.6μm laser signal by mixing with a reference signal and detecting the much lower frequency beat signal on a spectrum analyzer[27] (Figure 11). In that work, two lasers, offset from each other by a frequency difference Δ, were directed at an MIM diode through mirrors and lenses. A current source and instrument amplifier provided a dc bias current to the diode and a signal generator provided a constant ac current while two spectrum analyzers measured the low and high frequency response and a multimeter was used to measure the total AC and DC components of the diode response. The testing was set-up in this simultaneous manner because the diodes tested were of the point contact type and to keep consistency between different measurement methods the fragile diodes were hooked up to one testing set up where several different types of measurements could be made. The difference (beat) frequency method is a relatively simple way of verifying the very high frequency response of symmetric test diodes.

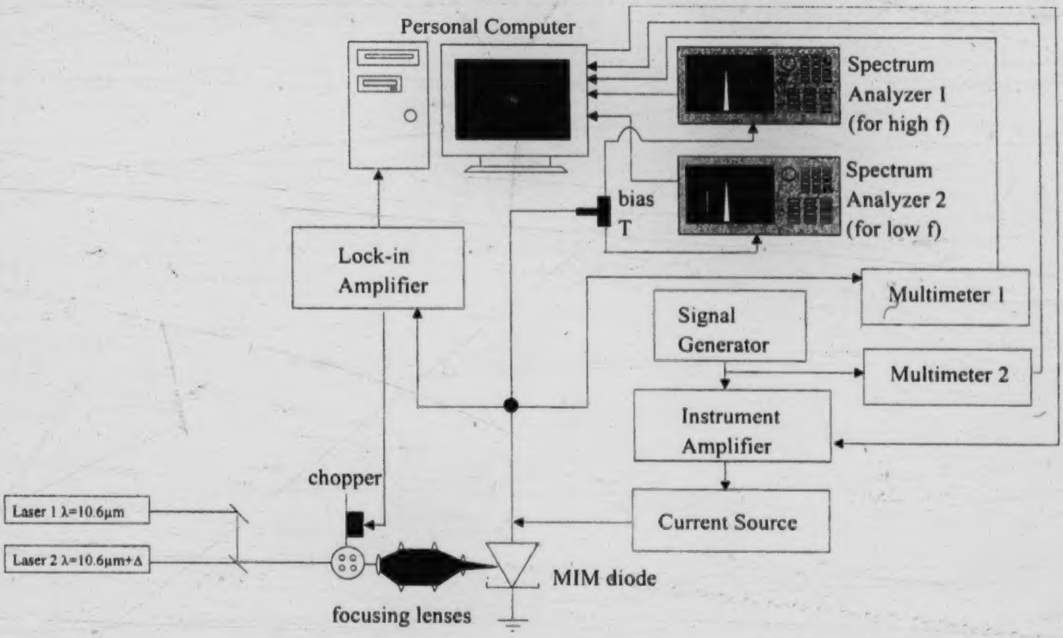

Personal Computer
Spectrum Analyzer 1 (for high f)
Spectrum Analyzer 2 (for low f)
bias T
Lock-in Amplifier
Multimeter 1
Signal Generator
Multimeter 2
Instrument Amplifier
chopper
Laser 1 λ=10.6μm
Laser 2 λ=10.6μm+Δ
Current Source
MIM diode
focusing lenses

Figure 11 Mixing of 10.6um lasers by an MIM diode (Masalmeh *et al.*[27])

3.4.3 Harmonic Radar

Since the MIM diode has a non-linear current-voltage characteristic it naturally produces harmonics of an applied sinusoidal signal (this effect can be illustrated by the fact that for a non-linear current-voltage characteristic the current can be described by an infinite polynomial series of order 2 or greater and substitution of a 'stimulus' sinusoid at frequency *f* into the polynomial leads to terms that have frequency *2f and/or 3f etc.*). This can be used to generate a 'response' signal in a harmonic radar system. Harmonic radar transceivers are simple devices that emit a harmonic frequency of an interrogating signal. This is advantageous over traditional radar in that the response signal inherently uses a different frequency so there is no confusion between the interrogating signal, signals reflected off of objects and the true tag response signal. Also, the energy of the interrogating signal is used to power the response signal, excluding the need for bulky batteries.

Harmonic generation by a diode stems from the fact that the diode has a non-linear *i-v* characteristic, with large second and/or third order terms when expressed by polynomial expansion. When a sine signal is substituted into a polynomial of second or third order it becomes readily apparent, after simple trigonometric identity substitution, that the resultant signal will have frequency components at two or three times the original applied signal frequency. The current of these frequency components can then be coupled to an output device, such as an antenna, to create a harmonic radar tag, which can be used to both receive and transmit. Thus, the application of the MIM diode to a harmonic radar system can be as simple as a diode connected to an antenna (Figure 12). In fact, harmonic radar systems used to track the movements of butterflies and other insects were composed of just a discrete Hewlett-Packard RF Schottky diode wire-bonded to a thin wire antenna and were detected by using a RECCO® transceiver.[42,43]

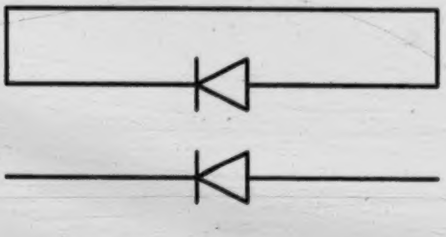


Figure 12 A simple harmonic radar device: folded dipole antenna (top) and single dipole antenna (bottom) tags

An optimal design may include a transmission line impedance transformer (the antenna impedance typically ranges from 2 ohms to a few hundred ohms while the dynamic resistance of thin film MIM diodes fabricated to date is typically in the tens of kilo ohms).

Harmonic radar systems have been used for radio frequency identification (RFID), tracking of small insects and animals,[44] vehicle toll collection, and most prominently, in avalanche rescue systems such as the one manufactured by RECCO. However, thus far none have been completely microfabricated, limiting the systems to simple (and relatively large) tags.

While the cost of a possible RFID system based on MIM devices may make them uncompetitive with many current low-cost RFID tags (25¢ and lower per tag with an ultimate goal of less than 5¢ per tag) MIM harmonic radar tags can fill several niche areas that are not filled by current devices. These niche areas generally consist of areas where the size of the tag is a critical concern. As a consequence of tag size, the frequencies of operation of the tag have to be high (to keep antenna size small). An ultimate goal is to fabricate a 'smart dust' system in which MIM diodes are mixing components in the R.F. transceiver circuit and rectification elements to power integrated MEMS sensors or other circuitry. This goal hinges almost entirely on MIM diode development since it can solve the two largest problems impeding the micro-miniaturization of tag transceiver systems: antenna and power supply size.

3.4.4 Rectification

If the electrodes in an MIM diode are made with metals of differing workfunctions the *i-v* characteristics will inherently be asymmetric. The turn-on voltages will be different for forward and reverse currents. The asymmetrical response will induce different forward and reverse currents in response to electromagnetic signals, such as that given by radio, microwave and possibly even solar radiation, producing an average D.C. offset current. This DC current can then be used as a power source for on-chip electronics. An MIM diode with an antenna and matching circuitry, known as a 'rectenna', has already been applied to directly detect a laser signal[8] and it has been proposed as a method of sunlight-electricity conversion.[45] The only limitations to how much radiation is converted to D.C. power are the efficiency of the antenna, the antenna to diode matching network and the asymmetry and non-linearity of the diode. As has recently been reported, lithographic metal antennas can be used as transducers of light to electrical energy.[46] A solar rectifier could then be as simple as a traditional RF antenna-diode rectifier scaled to visible wavelength dimensions.

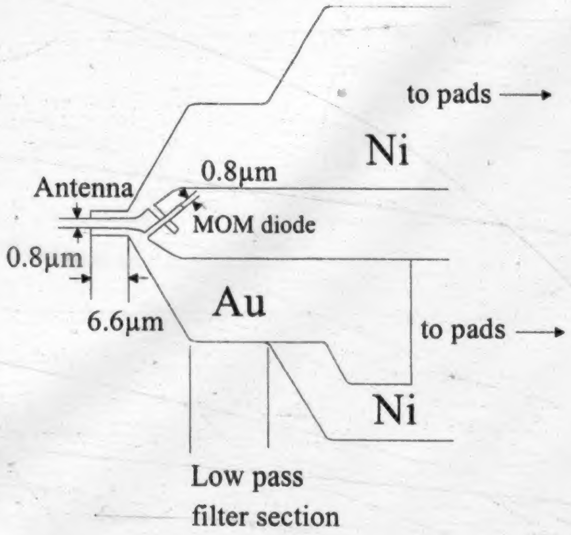


Figure 13 Rectenna used for the detection of laser signals (Hoofring *et al.*[8])

Despite its possible efficiency of greater than 80%[47] (compared to a maximum theoretical p-n solar cell efficiency of ~30% due to its frequency dependent operation) there has not been much research into the solar-power device application for the MIM diode. While a patent was granted on the concept of an antenna coupled to an MIM diode[48] there seems to be no evidence of a functioning prototype. Nevertheless, the application of the MIM diode as power conversion 'rectenna' seems particularly relevant for this research since *"current development indicates that MIM diode development is the main limiting factor that is preventing the optical rectenna from becoming a commercially viable next generation photovoltaic device.*[45]

Recently, through a contract from the National Renewable Energy Laboratory (NREL – Golden, CO, USA), ITN Energy Systems (Littleton, CO, USA) investigated the optical antenna-MIM diode system as a more efficient solar – electrical energy conversion device.[49] While they were able to demonstrate the concept through application of lower frequencies to high junction area devices, actual solar rectification was not achieved. One problem was that the higher junction area devices had good characteristics at high voltage (similar to the diodes produced in this research) but the good characteristics (i.e. high non-linearity) did not scale well with size and voltage (i.e. the nm size junctions which produced non-linearity at mV were only slightly non-linear).

While the ITN group did not completely prove the concept (partly due to time constraints, as this was government contracted work) the idea is still valid. Assuming an ultra small tunnel junction device can be fabricated with very non-linear and asymmetric *i-v* characteristics with a turn-on voltage (voltage at which the *i-v* characteristics transition from linear to non-linear) in the mV range (so that biasing is unnecessary), then coupling to optical antennas should yield rectified D.C. voltage, which can be added across many devices in a series-parallel array to output a large D.C. voltage and current.

3.4.5 Terahertz Imaging

A field that has been gaining interest recently is terahertz imaging (also known as T-ray imaging). With applications in medical imaging and diagnosis, security control, chemical and biological identification as well as internal defect detection, terahertz imaging has no shortage of possible applications despite its relatively early stage of development. Only in the past decade has research been significant, largely due to the availability of new devices and new funding from government sources such as the European Space Agency (ESA), the U.S. National Aeronautics and Space Administration (NASA) and the U.S. Defense Advanced Research Projects Agency (DARPA).

Different materials have different absorption and reflection responses to THz radiation. Non-polar molecule materials, such as fat, clothing and paper transmit THz radiation while polar materials absorb THz radiation with characteristic spectra.[50] This difference in absorption or emission characteristics can be used to advantage to instantaneously check for concealed weapons under clothing, for diagnosis of defects in structural material (such as foam insulation on the space shuttle), in medical imaging (unlike X-rays, THz waves are not harmful to tissue), or even for detection of dangerous biological hazards (most objects have a characteristic quiescent spectral absorption of THz radiation. Anthrax, for example, could be detected by analysis of its characteristic absorption spectrum).[51]

The THz frequency range (0.1-10THz) is a generally unexplored region of the electromagnetic spectrum, largely due to the fact that there are few suitable components for these frequencies.[52] At frequencies higher than the THz range, optical devices are used while at lower frequencies radio devices are used. The THz spectrum region is the frequency range where the optical (far-infrared) and electrical (GHz) spectrum meet, so it is an extreme region for both optical or electrical devices. The MIM diode is one of the few devices that *can* detect THz radiation; therefore it is ideal for application in THz spectroscopy. Further research is justified since the MIM diode has not yet been used for this purpose.

Research in THz imaging has begun to gain interest in the past few years. The European Space Agency started a program in 2002, named StarTiger, to create a real-time THz imaging device and reached its first goal shortly after by taking a still image of a concealed hand. A group at Delft University in the Netherlands[53] has used antenna-coupled bolometers to detect THz radiation through objects illuminated by a Ti:sapphire oscillator THz pulse. The antenna-coupled bolometers are THz responsive antennas on a heat sensitive material that is cooled to cryogenic temperatures to minimize heat 'noise.'

Two main techniques have emerged for THz imaging: THz time-domain spectroscopy (THz-TDS) and THz absorption imaging. The former uses the transient response of an object's materials to a femtosecond pulse, providing depth as well as material information while the latter uses simple amplitude data of an emitted signal passed through an object. THz time domain spectroscopy requires precise data acquisition and complex information processing of the transient waveforms. Absorption imaging relies on just a simple amplitude measurement of a transmitted or reflected signal.

The MIM diode is a good candidate for detection of signals for at least THz absorption imaging. To a first order approximation, an MIM THz transceiver could be as simple as an MIM diode fabricated with metals of differing workfunctions, a transmission line and a photolithographically patterned antenna. An array of these devices with feedhorn apertures providing individual pixel resolution and an addressing and gating scheme could comprise an image capture device.

3.4.6 Magnetic Tunnel Junctions

Recently, a lot of research and development has been done on magnetic devices and their practical applications, primarily as read/write heads for hard disk drives and magnetic RAM. These devices rely on one of three magnetic effects: anisotropic magnetoresistance (AMR), giant magnetoresistance (GMR) and spin-dependent tunnelling (SDT) (or tunnelling magnetoresistance (TMR)). Anisotropic magnetoresistance arises out of electron spin interaction with an applied magnetic field. It is generally a weak coupling between adjacent magnetic domains and thus produces a weak relationship between resistivity and magnetic polarization. Giant magnetoresistance arises out of spin-dependent scattering of electrons at ferromagnetic material interfaces. While GMR devices have larger magnetoresistances than AMR devices (~5-6% vs. 2-3%) these values are up to an order of magnitude less than those of spin dependent tunnelling devices. Spin-dependent tunnelling is a consequence of a relationship between electron spin and tunnelling probability. Specifically, the tunnelling probability varies due to tunnelling from up to up spin states and up to down spin states and vice versa.

Each of these effects demonstrates a fractional change in resistance $\Delta R/R$ with an applied magnetic field H which makes them ideal for application in magnetic field sensing and in non-volatile Random-Access Memory (RAM), since ferromagnets (FM) keep their polarization without applied power. Devices based on AMR and GMR (read/write heads) have already been in commercial production while SDT devices are still in the research phase (few, if any, commercial implementations of SDT devices currently exist). Nevertheless, SDT devices have received much more attention recently due to their high magnetoresistance (>30%) and high resistivity (low sense currents).

Magnetic Tunnel Junctions (MTJs) have been researched for more than twenty years but only recently has the sensitivity of the devices reached a significant level (magnetoresistance of 10% or more) at room temperature and hence the devices have only recently attracted more research interest.[54] The junctions are simply MIM diodes with the electrodes made of ferromagnetic material and are made with one electrode having a fixed magnetic polarization direction while the counter-electrode has an electrically alterable magnetic direction. The 'free' electrode is more susceptible to change its polarization with an applied magnetic field. This is done either by using hard and soft (low and high susceptibility materials) or by 'pinning' the fixed electrode by an adjacent polarized anti-ferromagnetic material (such as MnFe or NiO). In the process known as exchange biasing the magnetic moments of the anti-ferromagnetic (A-FM) and ferromagnetic material (FM) become coupled at their interface and the ferromagnetic material aligns itself throughout its thickness (assuming a thin layer) to its polarization at the A-FM/FM interface.

The tunnel barrier, typically Al_2O_3, is 10-30Å thick but the ferromagnetic and anti-ferromagnetic layers are also relatively thin (tens of Å) so that the magnetic properties are uniform across the layer. The tunnel current is a function of magnetic alignment of the two electrodes (i.e. when the magnetic fields of the two electrodes point in the same direction the resistance is minimum, when they are anti-parallel the resistance is maximum). This effect stems from the different density of states (DOS) distributions of up and down spin electrons in a ferromagnet. The electron current is split between up and down spins between the two electrode DOS distributions for each spin. When the two electrode polarizations are aligned the current consists of transfer between majority/majority spin and minority/minority spins across the barrier. If the electrodes are anti-parallel then current is from majority to minority DOS and vice versa, yielding a reduced conductivity.

Magnetic tunnel junctions have the advantage of higher sensitivity at lower magnetic fields over GMR and AMR devices (30% and greater vs. a few percent magnetoresistance each for GMR and AMR) and appear very much suitable for applications in MRAM, magnetic read/write heads and sensors. Compared to GMR and AMR devices the TMR devices have much higher impedances which allows for smaller sense currents to be used (and hence better power efficiency and faster access times). However, this higher impedance is detrimental to the signal-to-noise ratio since the Johnson RMS noise voltage is proportional to the square root of the device resistance.

Magnetic memory shows great promise to be the next generation non-volatile memory in computer systems, initially for 'instant-on' boot up and later as general-purpose non-volatile RAM. Its high speed, low power and unlimited read/write cycles have given rise to much research in MRAM (magnetoresistive random access memory) technology. While the processing challenges (namely layer uniformity and CMOS integration) have largely prevented current commercial production of MRAM chips it is not unreasonable to assume that such chips will be produced within the next couple of years.

Due to their higher magnetoresistance and lower sense current, magnetic tunnel junctions seem to posses the most potential, in terms of performance, for MRAM in comparison to AMR and GMR based devices. The architecture of a proposed design of a magnetic tunnel junction based MRAM is shown in Figure 14. The free layer position (parallel or anti-parallel) is sensed by the tunnel current across the junction, induced by the bias between bit lines (B1, B2...) and Word lines (W1, W2...) The current through the bit lines can then be sensed individually by amplifiers. The bit writing mechanism involves the data lines (D1, D2...) and bit lines. The free ferromagnetic layers are elongated, their long axis being the 'easy' axis or axis of high magnetic susceptibility. The short axis, or 'hard' axis is parallel to the data lines. The 'flipping' of the free layer involves the data and bit lines such that when current is passed through both, the memory bit at their intersection is set according to the polarity of the bit line current. The data line magnetic field serves to partially flip the bit towards the hard-axis of all junctions along its path but only the bit at the intersection with the bit line magnetic field is altered since it also has a magnetic field due to the bit line current in the direction of the easy axis.

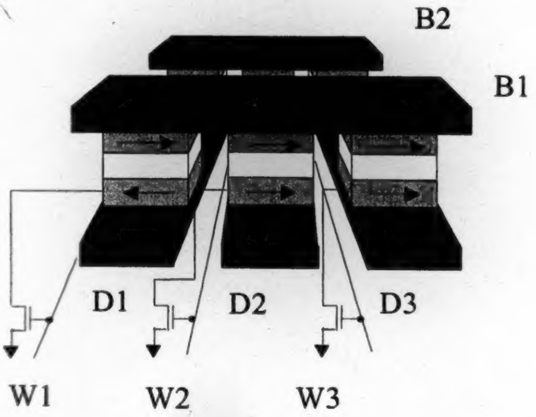


Figure 14 An MRAM 'cross-point' architecture with CMOS (*Tehrani et al* [55]), top ferromagnetic layer is free, bottom is fixed.

3.4.7 Superconducting Tunnel Junctions

An MIM diode constructed with superconducting electrodes and an insulator barrier is known as a superconducting tunnel junction (STJ) (also known as SIS junction or Josephson junction). Despite the requirement of cryogenically cooling, these junctions are state of the art components in high precision magnetic field sensing (medical magnetic imaging), low-noise mixing (radioastronomy) and X-ray detectors.

SIS Physics[56]

For the junction to be superconducting it must be cooled cryogenically to a temperature that is below the critical superconducting temperature of the electrodes. Above this temperature the electrons in the superconductor repel each other but below this temperature the electrons near the Fermi level of the superconductor are grouped into pairs, known as 'Cooper pairs'. These Cooper pairs consist of two electrons of opposite spin and behave as bosons. Consequently, all of the Cooper pairs can occupy the same quantum state. The Cooper pairs can tunnel through the insulating barrier and maintain phase coherence up to a critical current, I_0. Current above the critical current is due to single electrons. The current-voltage characteristic of a typical Josephson junction is given in Figure 15. As the bias current is increased the voltage across the junction remains 0 while the current is less than the critical current, I_0. Once the critical current is reached the voltage jumps to a non-zero value. An increase in current at this point results in ohmic behaviour while a decrease in current follows the curved portion of the trace.

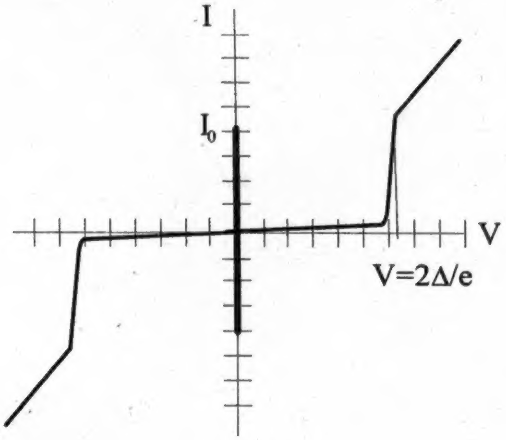


Figure 15 Typical *i-v* characteristics of a Josephson Junction (underdamped) with I_0, the critical current of the superconductors

A unique characteristic of the SIS junction is that there exists a current at 0 bias through a non-superconducting insulator. This Josephson current is due to Cooper pair tunnelling, which can be suppressed by exposure to a magnetic field of integer flux quantum. Within a superconductor all electrons, in the form of Cooper pairs, maintain the same wave-function phase. If the phases of the Cooper pairs are φ_1 and φ_2 in each of the respective superconducting electrodes then the current through a Josephson junction is given by:

$$J = J_c \sin(\varphi_1 - \varphi_2) \tag{24}$$

where J_c is:

$$J_c = \left(\frac{\pi\Delta_0}{2eR_J}\right)\frac{\Delta}{\Delta_0}\tanh\left(\frac{\Delta}{2k_BT}\right) \tag{25}$$

where T is the temperature and R_J is the resistance in the non-superconducting state and Δ is the superconductor gap. This relationship describes the DC Josephson effect. The AC Josephson effect is given by the relation:

$$\frac{d(\phi_2 - \phi_1)}{dt} = \frac{2e}{\hbar} V \qquad (26)$$

which, when integrated yields:

$$\varphi(t) = \frac{2e}{\hbar} Vt + \varphi_0 \qquad (27)$$

φ_0 is the magnetic flux quantum equal to $2.068 \cdot 10^{-15}$Wb. It can be seen from equation 26 that the current oscillates with frequency

$$2\pi f = \frac{2e}{\hbar} V \qquad (28)$$

and the junction converts DC voltage to AC frequency at a rate of 483.59767MHz/μV. This relationship is so precise and predictable that it has been used to most accurately define 1V (the definition of a volt through Josephson junction frequency to voltage conversion is known as the Josephson Volt Standard[57]). Further use of Josephson junctions for frequency generation includes the flux flow oscillator (FFO), which has been integrated with Josephson junction mixers in low noise receivers[58] and has been investigated for use as a clock oscillator for Josephson logic circuits.[59]

SQUIDs[60]

Josephson junctions are best known for their application in Superconducting Quantum Interference Devices (SQUIDs). These devices consist of a ring of superconductor, with its loop interrupted by one or more Josephson junctions (and a method to remove the hysteresis described earlier, such as shunting resistors in a configuration known as a resistive-capacitive shunted junction or RCSJ). The resulting structure is extremely sensitive to magnetic fields (as small as 10^{-14}T). This high sensitivity has led to their widespread implementation in measurement of the magnetic activity in the brain (of order 10^{-13}T), known as magnetoencephalography (MEG).

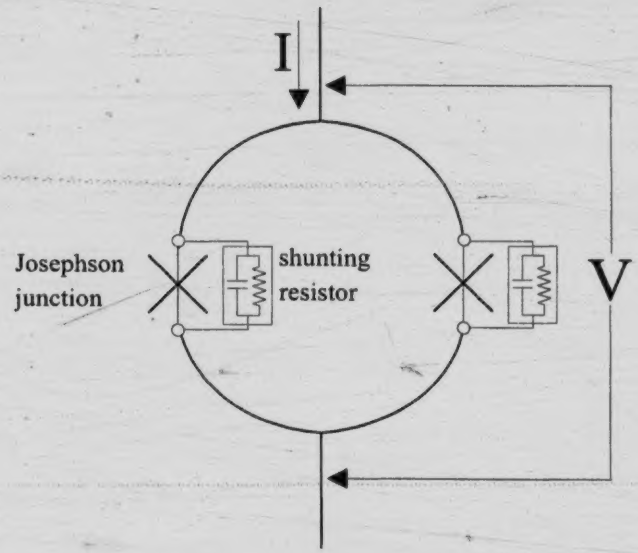


Figure 16 Schematic of a D.C. SQUID

In a D.C SQUID a D.C. current is injected into the loop and is split between the left and right current paths of the loop. A magnetic flux through the loop induces a circular current around the loop, which in turn induces a phase change between the left and right voltage components produced by the junctions. The resulting voltage across the loop oscillates with applied magnetic flux. Counting each voltage period corresponds to a change in flux of one quantum of magnetic flux.

A more complex, but historically easier to construct, device is the R.F. SQUID. The R.F. SQUID contains just one Josephson junction interrupting a superconducting loop. An LC tank circuit is inductively coupled to the superconducting loop and an R.F. current is applied to the LC circuit (Figure 17); the frequency is chosen such that the LC circuit is resonant.

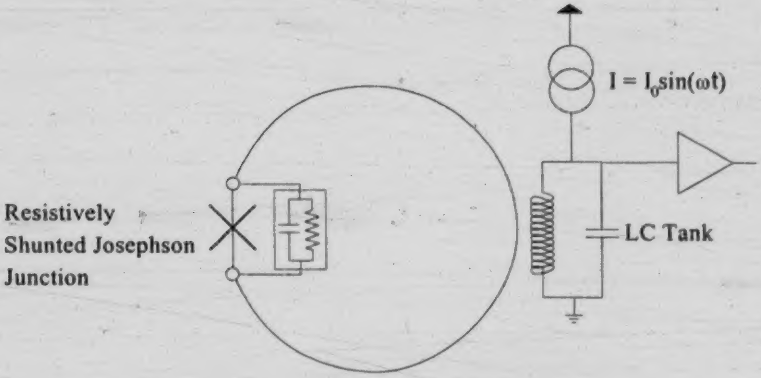


Figure 17 Schematic of an R.F.-SQUID

As in the D.C. SQUID, the R.F. SQUID outputs a periodic voltage with applied magnetic flux to the superconducting loop.

Fabrication of the superconducting loop and tunnel junctions is done using microfabrication techniques, either conventional photolithography or e-beam lithography methods. The loop itself is typically a square loop configuration with either a single loop or a spiral multi-loop. While historically the R.F. SQUID has been the more popular configuration due to its ease of fabrication (only one Josephson junction) the D.C. SQUID is now the state of the art standard. This is mainly due to advances in fabrication. The MEG systems in commercial production (by CTF Systems Inc. Vancouver, B.C. Canada and 4-D Neuroimaging, San Diego, Ca, USA) use D.C. SQUIDs with helium cooling (i.e. low temperature SIS junctions such as Nb/AlOx/Nb) and sense better than 10fT/Hz (sampling from D.C. to several kHz).

Low Noise Heterodyne Mixing

The heterodyne mixing properties of SIS junctions have been known since the late 1970s. Since then SIS junctions have become integral components in radioastronomy millimetre and sub-millimetre wave detectors due to their excellent noise properties (low shot noise due to low current) and low power demands on the local oscillator.[61]

Mixing in SIS junctions is due to quasiparticle (single electron) conduction as opposed to conduction by Cooper pairs, which are responsible for the zero resistance superconduction. Unlike Cooper pair conduction, quasiparticle current requires a bias voltage. In superconductors the density of states resembles the density of states of semiconductors more than metals. An energy gap exists between the filled and empty density of states (at 0K) that is equal to 2Δ. The Fermi energy lies exactly 1Δ above the top state of the filled density of states and 1Δ below the bottom state of the empty density of states. Conduction by quasi-particles takes place only when the bias voltage is greater than $2\Delta/e$ or at lower voltages if energy from photons makes up the energy shortfall. The gap voltage is typically a few mV.

The Cooper pair current (or Josephson current) is an undesirable source of noise in heterodyne mixers since it is not voltage dependent. To suppress the Josephson current an external magnetic field is applied to the junctions. The junction then follows a non-hysteric *i-v* curve, as in Figure 15 but with a zero current at 0V. An alternative method is to construct the barrier with a ferromagnetic material, such as NiO. This has been shown to expel the Josephson current while retaining good non-linear characteristics.

The theoretical upper frequency limit for mixing in an SIS junction is the frequency at which the photon voltage, $V_{photon} = \frac{h \cdot f}{e}$, equals twice the junction gap voltage $V_{gap} = \frac{2\Delta}{e}$:[62]

$$f_{\max} = \frac{4\Delta}{h} = \frac{2e \cdot Vgap}{h} = 0.484 \cdot V_{gap} \qquad (29)$$

where h is Planck's constant and $f_{\max}$ and V_{gap} are in THz and mV, respectively. For a superconducting material with a gap voltage of 2-3mV this yields a maximum frequency of 968GHz-1.452THz. Though matching circuit losses generally limit the frequency response to values significantly below the theoretical values the high frequency response of SIS junctions outperforms that of alternative mixers (such as GaAs Schottky diodes) at these frequencies, particularly in terms of receiver noise.

4

Chapter 4: Fabrication of Tunnel Junctions

4.1 Chapter Overview

The tunnel junction (or MIM diode) has been investigated for decades by many research groups for various applications, most notably as high frequency mixers. But despite the interest, it has only been recently that commercial applications of these devices have been actively pursued (CTF Systems, 4-D Neuroimaging). This has largely been due to the fact that fabrication methods have not been adequate to allow commercial production. Prior and current research has focused on the tunnel diode itself, and particularly the quality of the junction, rather than integration into new applications and more elaborate designs. This is indicative of the fact that, despite a long history of investigation, a high quality manufacturable MIM diode remains an elusive goal. This chapter presents fabrication details of thin-film metal-insulator-metal junctions published by researchers. The fabrication methods of conventional MIM, thick- insulator MIM-LCD, and superconducting SIS junctions are presented as a background to the current research.

4.2 MIM Junctions

MIM junctions have been constructed with thin film methods since the 1970s. However, due to the requirements of a very thin insulating layer, the devices produced have had varying characteristics and a definitive process that can be used to design MIM devices with a-priori known characteristics (apart from qualitative descriptions) does not yet exist. The only industrial uses of the MIM diode, thus far, have been in liquid crystal display panels, using a relatively thick, non-tunnelling junction, and in SQUID devices in (expensive) magnetoencephalography devices. Nevertheless, some success has been achieved in research and the various processes used to construct prior diodes can be used as a starting point for research, so that a cost-effective manufacturable process can now be developed and more types of devices (i.e. the MITT) become more feasible.

Since the critical process in the construction of the diodes is insulator quality, the insulator is invariably an oxide grown on an electrode metal. In general, deposited insulators have much higher surface roughness than grown oxides and the incomplete surface coverage (presence of pinholes) is a severe problem when dealing with layers of nanometre scale thicknesses. Also, metallic devices are cost-effective and, unlike other alternatives (such as heterojunction stacks) are much simpler to deposit and have therefore been the focus of most research into tunnelling devices. The choice of metals for the metal/metal oxide electrode is amongst those metals that can grow passivation surface oxides by simple exposure to oxygen or by exposure to oxygen plasma. These metals (aluminum, chromium, titanium, tantalum, nickel and niobium) have been used in tunnel junctions in various applications. For instance, tantalum grows a smooth and thick insulator (>10nm) therefore its use in high-field devices (such as LCDs (next section) is prevalent). Aluminum forms a stable oxide, Al_2O_3, but its large grain size and roughness, when deposited as a thin film, could hamper uniformity of device characteristics across a wafer.[63] Further, the native oxide grows to a thickness of 5nm in a very short period.[31] Any processing to a smaller oxide thickness would require that the sample not be removed from vacuum between the oxidation and top layer metal deposition steps. Chromium, does not naturally form a single stable oxide, at least when exposed to ambient air. Chromium forms three oxides (CrO_2, Cr_2O_3 and CrO_3) when exposed to ambient air, and possibly just Cr_2O_3 after an annealing step (at 300°C).[63] Cr/CrOx MIM diodes have had varying degrees of success in terms of non-linearity and not much is known about Cr based tunnel junctions. However, given that Cr is readily available, adheres well to substrates, has a fine-grain structure and has a native oxide of between 0.9 and 2nm (so that the lower limit on oxide growth is reasonably low) it is a good candidate material for tunnel junction fabrication. Titanium is another metal that grows a native oxide that is suitable material for tunnel junction fabrication. The native oxide formed is largely a function of the oxygen partial pressure though TiO and TiO_2 are the main oxides formed with Ti_2O_3 and Ti_3O_5 also formed to some degree.[64] Given this, titanium may not be the best choice of metal for tunnel junction fabrication, until a suitable process can be found to grow a stoichiometric oxide on the titanium surface. Niobium, typically used in superconducting junctions, does not form a stable oxide[65], but can be made to be somewhat stable by using a non-reactive metal as the counter electrode. It has also been reported that using a plasma oxidation technique yields a stable enough oxide to produce all-niobium junctions.[75] Nevertheless, niobium does not seem as a safe choice for initial MIM diode fabrication. Finally, nickel is a material commonly found in literature when discussing point-contact and even thin-film room-temperature

MIM diodes. The choice of nickel has largely been heuristic, but the results achieved with nickel have surpassed that of any other metal combination tried.[7,28] It is known that Ni forms a single stable oxide, NiO, and that the oxide forms naturally to a thickness of approximately 1nm.[66] Given this, and the fact that nickel has been the most successful metal/metal oxide electrode combination, it seems the obvious choice with which to begin constructing tunnel junctions.

The substrate used for fabrication of MIM diodes was most commonly SiO_2, thermally grown on conventional Si wafers. The substrate does not generally have a great effect on device characteristics at low frequencies, but the substrate properties must be accounted for when the MIM diode is used as a high-frequency device (mixer, rectifier etc…). Glass and sapphire substrates have also been used, but given that the SiO_2 covered Si wafers are readily available, nearly atomically flat, readily oxidized, inexpensive, inherently compatible with the microfabrication tools and can be diced, they are the most obvious choice for MIM diode processing and may also be a desirable choice for harmonic radar tag development.

The most common method of patterning the electrodes has been single-layer lift-off lithography. Unlike wet and dry etching techniques this method minimizes the possibility of chemical interaction with the electrode surface. Furthermore, any metal can be patterned in this manner. Typically, a reverse bevel profile is desired at the photoresist sidewall. The reverse-bevel (or overhang) ensures that metal on top of the photoresist does not make physical contact with the photoresist on the substrate, at the bottom of the photoresist channel. This ensures a clean lift-off once the photoresist is dissolved in acetone. The overhang is typically achieved by 'hardening' the top of the photoresist in chlorobenzene or toluene prior to development. The process is shown in Figure 18.

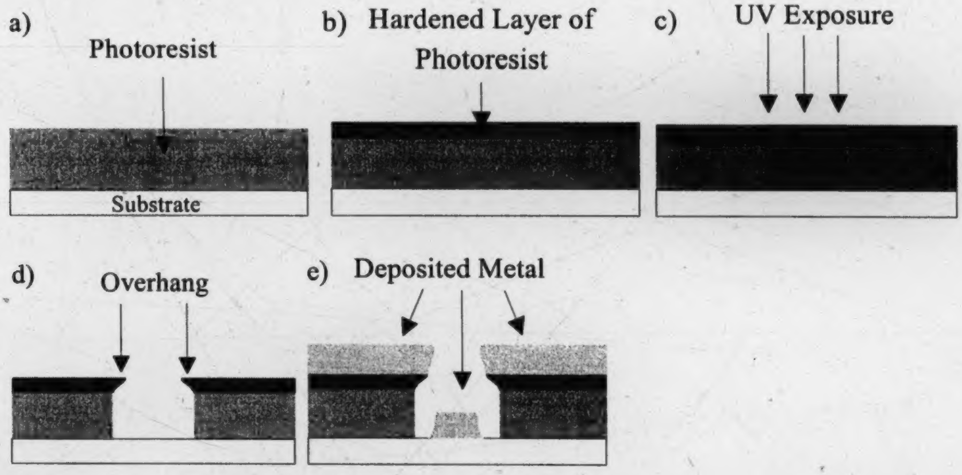


Figure 18 A lift-off process a) photoresist is spun on a wafer b) the surface of the photoresist is depleted of photoactive compounds by soaking in chlorobenzene c) the sample is patterned with a photomask and UV light d) an overhang results on the sidewall d) metal is deposited.

Oxide growth has been the most critical of process steps in the development of MIM junctions and several methods have been used to grow the oxide on the metal electrode. Historically, the metal was simply allowed to oxidize naturally in ambient air. This method inherently has poor control over insulator thickness, as the thickness is determined by metal properties and adsorbed impurities. A plasma oxidation technique, wherein the electrode is exposed to a low-power R.F. oxygen plasma for a timed period, has given more control and more reliable junctions.[26] The process used by Havemann[26] to fabricate edge MIM devices included a sputter etch of the nickel surface in Ar^+ plasma (5mtorr at 0.137 W/cm^2 for 2 minutes) and an oxidation in 5-30 mtorr oxygen plasma at a power of just 0.055 W/cm^2 for 2 to 15 minutes. The resultant devices had oxides of 1-2 nm in thickness. A more versatile process, presented by Greiner for the fabrication of Josephson junctions, was the Ar^+:O_2 sputter-etch mode oxidation.[101] Greiner used his process to fabricate cross-electrode Pb(In)-Pb oxide-Pb and Nb-Nb oxide-Pb junctions with estimated oxide thicknesses ranging from 20Å to 50Å while achieving a high number of functioning devices (over 90%). The process involves exposure of the sample to plasma of Ar and O_2. The ratio of Ar^+ to O_2, total pressure and R.F. power determine the thickness of the resulting oxide grown. This is advantageous over pure oxygen plasma or ambient exposure since it allows control over the oxide thickness without a time dependence. The Ar^+ plasma etches away surface atoms, such as any oxide on the metal surface, while the oxygen plasma provides energetic free radicals for oxidation of the metal surface. The expression for oxide thickness x as a function of time, initial oxide thickness and sputter and deposition rates R and K was found by Greiner to be:

$$x = x_0 \ln\left(\frac{K}{R} - \left(\frac{K}{R} - e^{x_i/x_0}\right) e^{-Rt/x_0}\right) \tag{30}$$

K and *R* are oxidation parameters dependent on power, temperature and pressure.

The two processes, sputtering of surface oxide and growth of oxide, eventually reach steady state such that the rate of oxidation is equal to the rate of sputtering:

$$\left(\frac{dx}{dt}\right)_{oxidation} = Ke^{-x/x_0} = \left(\frac{dx}{dt}\right)_{sputtering} = R \tag{31}$$

The competing processes of Ar^+ sputter etching and oxidation reach a steady state after about 5 minutes (in the present research) reliably yielding an expected thickness of metal oxide.

In general, full process details for MIM junction research have not been made available in publications. The details of one process for fabricating edge and cross-layer Ni-NiO-Ni diodes were published by Heiblum *et al* [28] and are presented here. For high-frequency experiments (0.6328μm and 3.39μm wavelength) a quartz substrate was used by Heiblum's group, while an n-type Si wafer with a 500Å SiO_2 isolation layer was sufficient for lower frequency (10.6μm wavelength) experiments (but still lossy at these frequencies). A 100Å layer of Ni and a 2000Å layer of SiO_2 were sequentially deposited in a multi-target R.F. sputtering chamber (Perkin-Elmer NRC 836: 5" target, 2.5". target substrate separation) without breaking vacuum between depositions. For the cross-layer diodes the deposition of the SiO_2 over-layer was omitted. The wafer was then removed from the sputtering chamber and photoresist was spun and patterned for the first electrode (the metal not covered by photoresist was to be etched away). The photoresist was then baked for 2 hours with a gradually increasing temperature up to 160°C. The wafer was then sputter-etched (presumably in Ar^+ plasma), which removed the exposed SiO_2 and Ni layers, as well as 500Å of the SiO_2 substrate. The photoresist was then removed in an oxygen plasma and the wafer was reintroduced into the sputtering chamber. After a base pressure of 10^{-7} Torr was reached, argon was introduced into the chamber and the native oxide on the Ni electrode was removed with an Ar^+ sputter etch. Re-oxidation of the Ni electrode then followed by introducing oxygen into the argon filled chamber in a ratio of 4:1 $Ar:O_2$ at a total pressure of 5-20mtorr and applying a D.C. voltage of 80-100V for 10 min. The system was then evacuated and a 1000Å layer of nickel was sputter deposited. The diode was completed by patterning the wafer with photoresist and chemical etching of the top nickel layer (the bottom electrode was insulated from the etchant by the grown NiO).

4.3 LCD

MIM diodes have already been used in industry, most notably as components of a liquid crystal display (LCD). Proposed in 1980 by researchers at Bell Northern Research[67] MIM-LCDs were the first active matrix displays. While they have since been replaced by thin-film transistor LCD panel technology, MIM LCDs are still in production in low cost displays where low power and small pixel size are competitive advantages. Figure 19 shows a cross sectional view and schematic representations of a typical MIM LCD pixel element. The pixel diode functions by acting as a threshold control device; allowing a large current to pass when the pixel is addressed with an applied voltage across the row/column lines and little current passes when the pixel is not addressed. The voltage that is built up across the liquid crystal is sufficient to keep the liquid crystal 'on' for the rest of the frame.

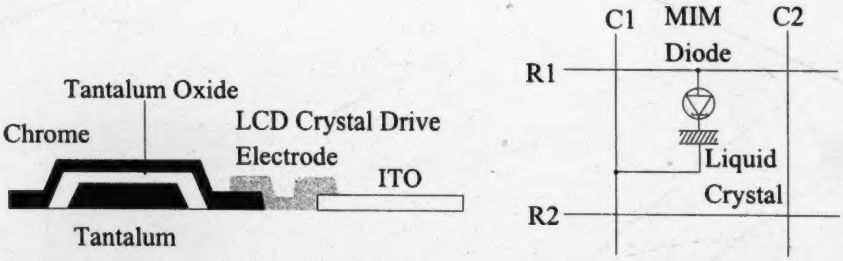


Figure 19 MIM-LCD of Chrome, Tantalum Oxide and Tantalum (left) and addressing schematic of an MIM pixel (right)

The fabrication process for the MIM-LCD is relatively simple. Typically, a 100-300nm D.C. sputtered tantalum electrode is patterned by dry etching (CF_4/O_2 plasma or some other fluorine containing plasma). It is then oxidized in an acidic solution (such as 0.01wt% citric acid[68,69] or 0.01-1wt% ammonium borate[70,71]) by application of an anodizing voltage (~30-50V) to the electrode for about 1h. The resulting insulator (TaOx) is grown to a thickness of 20-70nm, depending on the anodizing voltage, on the surface of the tantalum electrode. This process yields a relatively smooth oxide on the tantalum electrode with few defects, however an annealing step is sometimes used to modify the barrier stoichiometry (from Ta_2O_5 to TaOx when annealed in vacuum to back to Ta_2O_5 in air) thereby modifying device parameters, such as turn on voltage (which increases with increasing annealing temperature).[69] The counter-electrode (chromium is typically used as it yields a diode with a more symmetrical *i-v* characteristic) is then deposited by CVD or PVD and the diode is complete.

By making the insulator thickness relatively large (in comparison to traditional MIM diodes which have insulator thicknesses on the order of ~1-10nm) the designers solved the problem of MIM diode reproducibility and reliability. At thicknesses of 20-70nm pinholes and hot spots are not concerns as the surface roughness of the oxide/metal interface is at least an order of magnitude smaller than the average thickness of the oxide. Also, the statistical variation of insulator parameters (i.e. effective thickness, defect density) between devices on a wafer is minimized as the thickness is increased. However, this solution does have major drawbacks when considering the device for other applications; the first of which concerns the driving voltage. To achieve a suitable 'on' current with low 'off' state current the voltage across the device must typically be 10-20V, far too high to apply to devices in other applications, such as in R.F. front-end systems or logic circuits. Furthermore, the process to grow the insulator requires that the devices be connected in parallel for parallel fabrication to be possible (i.e. a wafer of independent devices would require a separate application of anodic voltage to each device). Also, the conduction mechanism through the oxide is Poole-Frenkel conduction (conduction through imperfections in the oxide) and not strictly electron tunnelling, so that the ultimate frequency response is much lower than is possible with tunnelling devices. The MIM-LCD solution to thin insulator problems is thus not suitable for most of the applications presented here.

4.4 SIS Junctions

Construction of SIS junctions of low-temperature superconducting material has been varied but certain materials and techniques have been repeatedly used due to their performance. Apart from a wide variety of high temperature superconductor materials SIS junctions have been constructed mainly from niobium and lead alloys. Due to poor thermal cycling properties (from room temperature to a few K and back) lead has not been an optimal choice for SIS junctions; nevertheless lead alloys with better thermal cycling properties have been used extensively in research. Niobium is the highest temperature elemental superconductor (T_c = 9K), so its use in SIS junctions has been extensive. Initially, in the search for an all refractory metal SIS junction, niobium was used in both electrodes and the oxide barrier.[72,73] Junctions made only with niobium have not given good characteristics, due to the non-stoichiometric niobium oxide being reduced by the niobium during deposition of the second electrode.[33] To circumvent this problem the second electrode has been constructed using other materials, such as lead alloys or the high temperature superconductor YBCO. It was also found that when an R.F. plasma oxidation technique was used this problem was not prohibitive and all-Nb junctions could be fabricated.[75] Alternatively, a higher quality junction is now typically made with aluminum oxide (which naturally forms in only one stoichiometry) as the insulating medium. The aluminum oxide barrier is either deposited by sputtering[74] or an aluminum layer is deposited (typically less than 10nm thick[33,36]) then oxidized with a thermal[33,36] or plasma[75] oxidation technique. Typical base pressures for both the niobium and aluminum are in the $1\text{-}10 \cdot 10^{-7}$ torr range[65] (deposition rates unspecified) while the deposition pressure in sputtering is usually 1-10mtorr and the niobium thickness ranges from about 10nm to over 200nm, The resulting structure is a $Nb/Al/Al_2O_3/Nb$ multilayer SIS junction (the aluminum is thin enough that it is superconducting by proximity to the niobium). This multilayer has seen much research and has gained wide acceptance as the configuration of choice in low temperature SIS junctions.

There are several geometric configurations of SIS junctions. The most widely used is the vertical stack, which is most commonly deposited at once, then patterned with dry etching. The Nb layers are typically defined by etching in an RIE chamber using a 90% CF_4/10% O_2 mixture (a process that can be readily implemented at the University of Alberta Nanofab) in a process known as Selective Niobium Etching Process (SNEP).[76] The tunnel junction itself is made through a via in an SiO_2 layer that separates the two electrodes. A version of this process was used by researchers at the University of Alberta and the Alberta Microelectronic Corporation (AMC, now Micralyne, Edmonton, AB, Canada) to construct NbN/MgO/NbN tunnel junctions[77] and Nb/Al/AlOx/Nb tunnel junctions.[78]

The NbN/MgO/NbN tunnel junctions made at the University of Alberta were constructed entirely with sputtered material except for an evaporated SiO passivation layer (D.C. sputtering of Nb in an Ar^+/N_2 plasma for the NbN deposition and an R.F. sputtering of MgO from an MgO target in an Ar^+ plasma). Their devices achieved good characteristics, which were comparable to those of other researchers, but had high sub-gap leakage current, which they attributed to poor quality of the MgO layer. A TEM image of their device showed a rough SIS interface, which may have had a role in the poor MgO layer. However, the direct deposition of MgO on NbN has produced problems for other researchers as well.[79]

The Nb/Al/AlOx/Nb tunnel junctions made by the AMC for a University of Alberta researcher (Veidt[78]) had also exhibited problems associated with film roughness. Aluminum was deposited to a thickness of 13nm immediately following the Nb sputtering (D.C sputtering in a load locked Perkin-Elmer 4400 Delta with a base pressure of 10^{-7} torr) in the same system without breaking vacuum. The SiO deposition was done in a thermal evaporator with a base pressure of 10^{-6} torr, then etched away at desired junction sites. The method of Al oxidation was exposure of the deposited Al layer to a load lock filled with O_2 at 60 mTorr for 3 minutes. The top layer of niobium was then deposited to a thickness of 90 nm followed by a 30 nm layer of gold (to prevent oxidation of the niobium). The patterning was done using lift-off and RIE. Their final structure is depicted in Figure 20.

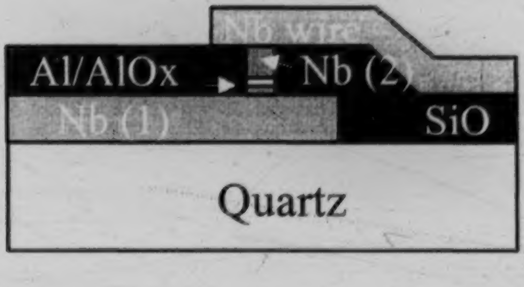


Figure 20 The Nb/Al/AlOx/Nb structure created by Veidt *et al*[78]

During the manufacture of their SIS junctions by the AMC, Veidt *et al*, found the oxide quality to be lacking in early runs in that pinholes seemed to be a major problem (high leakage currents were found). The group investigated the problem and found that as the niobium thickness was increased, the interface between the aluminum and niobium became less sharp. The problem was deemed to be surface roughness of the niobium, which increased with thickness. The only solution that was feasible was to increase the thickness of the aluminum. Their other (unsuccessful) attempts at making a smoother film by decreasing the deposition pressure (which produced thin films with high internal stresses) or their idea of depositing a thinner layer of metal (which they rejected since it was required that the layer was thicker than the magnetic penetration depth of the metal) also seem to be viable solutions to consider for tunnel junction fabrication with other metals since those negative characteristics can either be dealt with or have no significance in this research.

5

Chapter 5: MIM Diode Process Development

5.1 Chapter Overview

This chapter presents the fabrication details and characterization of the MIM tunnel junctions of the present work.

5.2 MIM Design

There were two main methods of fabrication for the MIM devices considered for this research: the cross-layer and edge-junction. The cross-layer could be made entirely with sputtering and without etch-back steps, making it a more accessible fabrication method, while the edge-junction method, made at least partially with e-beam evaporation, is a lower contact area device, and therefore higher frequencies are possible. The process for the cross-layer junction is given in the next section while the process for the edge-junction is given in Appendix B (for machinery specification see Appendix A).

The diodes on the 'MIMS' design wafer (see Appendix C) were arranged into rows of devices of geometrically identical dimensions. There are 7 columns of devices and 13 devices in each column. The 1st column, closest to the centre of the wafer, consisted of 'shorted' devices that were simply devices that had no tunnelling contact, but rather had a metal – metal contact, to be used as a reference. The remaining columns 2-7 had devices of varying contact area, defined by the varying line-width of the top electrode

(1μm, 1.5μm, 2μm, 3μm, 5μm and 10μm). The diodes were referred to individually by their column number and an alphabetic letter (A-N) to reference their position from the 'top' of the wafer (i.e. diode 7B is in the 7th column, 2nd device from the top). See Appendix C for wafer map.

Each wafer included three test areas to more adequately characterize the process. The first test area is a 7.5mm by 15mm rectangle that is used as the contact area for 4-point probe measurements as well as for ellipsometric measurements. A second area consisted of two large pads and a connecting thin line. The idea is to test the adhesion between two layers. The pads are fabricated by metal deposition in one step while the thin line is deposited later. The final test area consisted of three long lines of widths 1μm, 2μm and 3μm, each of length 4.665mm. These lines were all patterned on the same mask as the second electrode such that a resistivity measurement can be attained for the metal deposited on this layer, since the 4-point probe area was constructed for the MIMSLF and MIMSBOE layers only (described below), to save on mask-writing time.

The MIMS process is a three mask process (see Appendix C) wherein the bottom electrode is patterned first mask (MIMSLF), followed by the etch mask (MIMSBOE) to etch back any oxide deposited on the first electrode, and finally the second electrode mask (MIMSSF) is used to pattern the top electrodes. The masks were written in-house by a Heidelberg[80] DWL-66 Pattern Generator using 40mm and 4mm lenses. The pattern was designed using L-Edit Student Version, a commercial software program for semiconductor and MEMS layout, and was converted to the GDSII format using the full version of L-Edit[81] and subsequently converted to the pattern generator's proprietary format using the Convert program (Heidelberg Instruments), made available at the

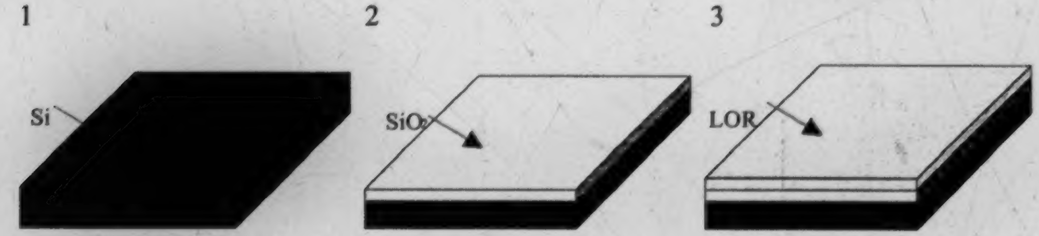
University of Alberta Nanofab. The first two masks, MIMSLF and MIMSBOE, were written using the 40mm lens which can produce mask features as small as 10μm while the last mask, MIMSSF, was written using the 4mm lens which, while it takes a significantly longer amount of time to write (and therefore more expensive), can produce features as small as 1μm. Also for reasons of expense, the antenna devices were not written (finished) on the small features mask (for a 70% reduction in cost from $770 to $230).

5.3 Cross-Layer MIM Diode Process

The MIM diode processes were developed first by investigation into previous researchers' methods. A survey of background research showed that lift-off was the most common method of patterning for this device (due to a lower probability of contamination than with chemical methods) and that the devices were often constructed on SiO_2 covered Si wafers (an inexpensive insulating substrate). A lift-off method using two resist layers was chosen as a more effective alternative to the traditional method of chemical treatment of the photoresist using chlorobenzene, which was found to not be compatible with the photoresist used (see Section 5.4.3). It was intended to deposit all device material using e-beam evaporation but due to problems with adhesion using this machine and its availability sputtering was used (both problems discussed in the Physical Characterization section of this chapter)..

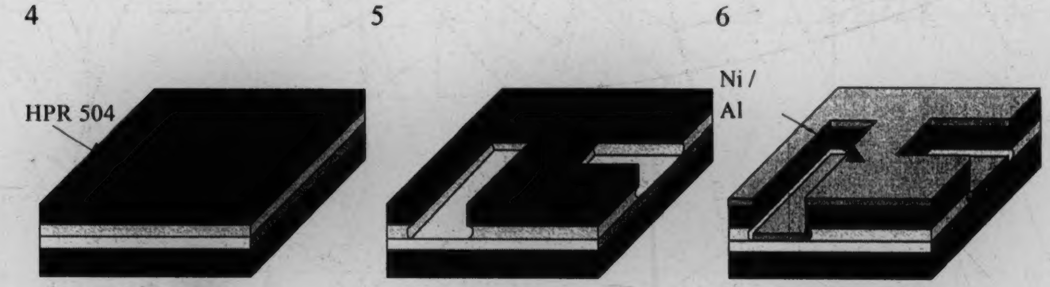
In each deposition of Ni an Al adhesion layer was used, as Ni does not adhere well to SiO_2, while Al does. For this reason, at least one group has used Al as an adhesion layer for Ni in the fabrication of MIM diodes.[8] Ti and Cr were also considered as adhesion layers but were not implemented in an effort to keep the number of variables to a minimum (oxidation properties in relation to Ni are unknown, whereas Al is known to naturally oxidize more rapidly than Ni). However, in the future, better adhesion may be attained by using these metals.

Process Steps:



1. A bare silicon wafer is first soaked in hot Piranha ($3H_2SO_4$:$1H_2O_2$) for 30 minutes to remove contamination. It is then rinsed with de-ionized (D.I.) water in a drop-deck for one 5-minute cycle and then rinsed and dried in a spin-rinse dryer for a 3-minute cycle (1 minute rinse and 2 minutes spin-dry).
2. An insulating substrate is required for the devices. Rather than use an expensive sapphire substrate or relatively rough glass, SiO_2 is grown on a silicon wafer through wet thermal oxidation. The p-type 4" wafers are placed in a furnace with a nitrogen bubbler for two intervals, between which they are turned 180° and re-oriented front to back for better uniformity. Typical parameters are 2.5 hours intervals at 1000°C to yield a layer of ~800nm of SiO_2 ("Filmetrics"[82] spectral reflectance measurements). The thickness is a non-critical parameter but to decrease capacitance thicker is better so the time in the furnace and temperature can be increased.

3. After a dehydration bake of 5-10 minutes on a 200°C hotplate, adhesion promoter hexamethyldisilazane (HMDS Arch Chemicals[83] Part #844074) is spun-on at 4500r.p.m. for 60s onto the wafer followed by a further 5 minute bake on a 150°C hotplate. The wafer is then coated with Microchem's Lift-off resist 5B[84] (LOR 5B) (dispensed at 500 rpm for 5s and spun at 3000 rpm for 45s) and pre-baked on a hotplate at 150°C (measured by infrared thermometer) for 5 minutes.



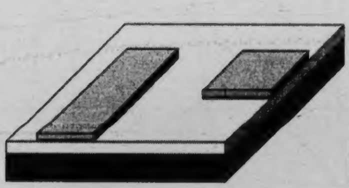
4. Positive photoresist (OCG HPR-504[85], Nanofab standard resist) is spun at a rate of 5000 rpm to yield an HPR 504 photoresist thickness of 1.2μm. The photoresist is pre-baked on a hot-plate for 60 seconds at 105-110°C and cooled for 1 minute at room temp.

5. The photoresist is exposed to mask 1 (large features) for 4s using either the "Elmo" or "Oscar" mask aligners (ABM Contact Mode with Zeiss optical alignment - "Elmo" or CCD camera alignment - "Oscar". Elmo was used when possible for consistency, but this was not always possible as Elmo was converted to a deep UV system for other projects for about three months). The photoresist is developed with MF-354 developer for 10-20s. The wafer is then taken back to the vacuum hotplate and hard-baked at 120°C for 2 minutes. The LOR is etched for an undercut with a tetramethylammonium hydroxide (TMAH) based developer

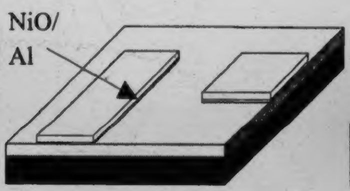
(MF-321) for 1 minute (yielding an undercut of approximately 5μm). Due to the hard-bake the HPR-504 photoresist should not be etched significantly by the TMAH based developer. The wafer is then flooded for 30s (UV exposed again without a mask to increase photoresist solubility for easier removal in step 7)

6. The wafer is taken to the sputtering chamber ("Doug") with Ni and Al targets loaded (K.J, Lesker[86] 99.995% pure 3", 0.125"thick). The chamber is pumped down to a base pressure of 10^{-8}torr and the process gas (Ar) is introduced and the pump gate is opened to an amount such that an equilibrium working pressure of 0.7mtorr is reached. The Al and Ni targets are separated by a cross-contamination shield and the Al targets are burned in with their shutters closed for 5 min at 200W immediately before being used. The Al target is used first to deposit a thin layer of Aluminum (1 minute of exposure with the wafer-shutter open and a target-sample 45°/distance of ~10cm) then the gun is turned off and nickel burned in on the second gun for a duration of 5 minutes. The nickel gun is also set to 200W with the same target-sample distance and angle and the shutter is opened (after the burn in) and deposition runs for 5 minutes to yield approximately 100nm of Ni film (measured by profilometer after deposition).

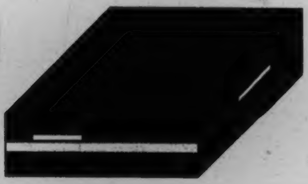
7



8



9

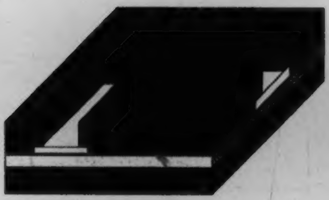


7. The photoresist is dissolved in acetone until all undesired metal is lifted off (with no agitation). The LOR resist is dissolved with LOR microstrip (Microchem

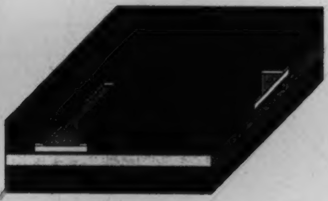
Remover PG). The wafer is cleaned with DI water for a 5-minute cycle in the dump-rinse and a 3 minute cycle in the spin-rinse dryer.

8. The wafer is taken to a reactive ion etcher (RIE) that has been thoroughly cleaned by wiping down with isopropyl alcohol (IPA) and then running a 100% O_2 200W plasma at 0.1 torr for a minimum of 30 minutes. Oxygen flow and the R.F. power is then turned off and the remnant oxygen is evacuated. The chamber is then preconditioned with 100% Ar plasma at 100W at 0.1 torr for 5 minutes. The chamber is then considered suitably clean and the wafer is placed in the chamber and cleaned of native oxide with an Ar plasma at 100W for 10-15 minutes. A layer of NiO is grown by exposing the wafer to an Ar:O_2 ratio and power to yield a desired NiO thickness (i.e. 95:5 Ar:O_2 plasma at 45W and 150mtorr for a time exceeding 5 minutes yields a 30Å thick NiO layer – as discussed in Chapter 5:)
9. HPR 504 Photoresist is spun on as in step 4.

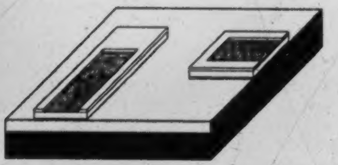
10



11



12

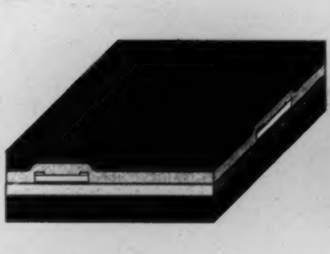


10. The photoresist (no LOR this time) is patterned with the pads mask (MIMSBOE mask aligned with alignment marks on mask and deposited features on wafer to +/- 1μm) and developed as before (or with the usual NaOH)
11. The wafer is then exposed to an ammonium hydroxide solution to remove the NiO from the exposed areas then rinsed with IPA and DI water. (Adapted from CRC

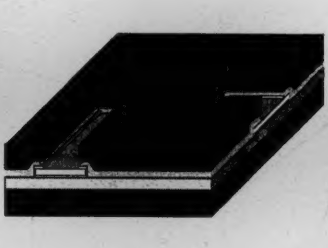
Handbook[87]) This step enhances the probability of an ohmic contact between the bottom layer of Ni and a subsequent Ni deposition onto exposed (non-junction) areas. Note: Alternatively, a high power Ar$^+$ plasma can be used to remove the NiO from exposed regions or, for non-integrated diodes, this step can be omitted as contact pressure from a probe during testing of a finished device virtually assures ohmic contact between this metal layer and a subsequent deposition layer.

12. Photoresist is stripped as before with acetone and cleaned in the dump-rinse and spin-rinse dryer for one 3 minute cycle each.

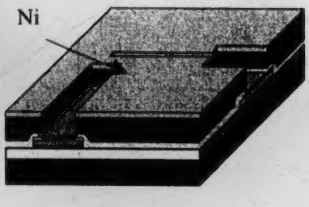
13



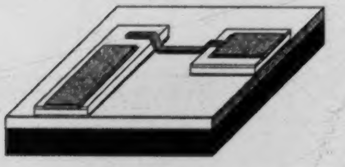
14



15



16



13. LOR 5B and HPR 504 photoresist are re-spun as in steps 3 and 4.
14. The wafer is manually aligned with the patterned alignment marks and exposed for 4 seconds each to the second electrode and pad windows masks (Masks exposed successively). The wafer is developed as before with MF-354, hard-baked and developed again with MF-321 or CD-26 for 15s.
15. Layers of aluminum and nickel are sputtered on the wafer (300W at 0.7mtorr Ar$^+$ pressure after pumpdown to 10^{-8} torr) to thicknesses of approximately 10nm and

200nm respectively. For better step coverage the deposition pressure can be higher than the 0.7mtorr. Initially lower power is used until a sufficient thickness is achieved so as not to damage the thin NiO layer by ion bombardment. E-beam evaporation is also used as an alternative to sputtering the second electrode (base pressure ~1-2 10^{-6}). Both methods were used in this research.

16. The photoresist is removed as before. The wafer is cleaned in the dump-rinse and spin rinse dryer for one 3-minute cycle each.

5.4 Physical Characterization

5.4.1 Section Overview

Physical characterization was carried out throughout each process to quantify variables in search of optimum process results and as checks for consistency. Both the device materials and process steps were characterized, the metal through resistivity and roughness measurements, the oxide through ellipsometric measurements and the process through inspection by SEM. This chapter presents the methods of characterization and their results.

Prior to the presentation of the characterization of the process it should be noted that the fabrication took place in a multi-user facility where each piece of equipment had multiple uses and users. As a result, the equipment (particularly the RIE and e-beam evaporator) was usually contaminated with a complex mix of various process gases or evaporants (visible film of adsorbants on chamber walls before cleaning and, in the case of the RIE, the noticeable smell emanating from the open chamber). Despite extensive efforts to clean the equipment prior to use it was accepted that each process step would have associated with it a certain level of uncertainty due to the cross-contamination between various users' processes. Further, some of the equipment was in the process of being commissioned, was being repaired or was intended for other applications. The oxidation process, for instance, is not a 'standard' process for the RIE, in that it is a relatively unique use of this equipment. As a result, the process was largely experimental in nature.

5.4.2 Metal Characterization

The first device material characterization tests were resistivity measurements of the sputtered and e-beam evaporated thin films. A 4-point probe was used to measure the resistivity of a 7.5mm x 15mm rectangle approx 150nm in thickness (measured by mechanical profilometer) present on each wafer of devices as well as an unpatterned SiO_2 wafer covered with a film of e-beam evaporated or sputtered film. The four-point probe accurately measures the resistivity of thin films by probing the voltage drop across a region of known current in the film (Figure 21). Measuring resistivity in this manner removes the probe resistivity from the measurement.

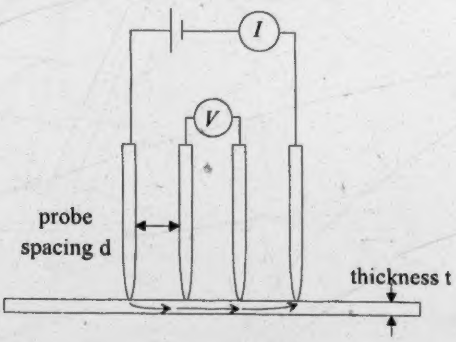


Figure 21 Four point probing of a thin sample

The relationship between the measured value (V / I) and the film resistivity is given by

$$\rho = R_s \cdot t = k \cdot \frac{V}{I} \cdot t \qquad (32)$$

Where ρ is the resistivity, t is the film thickness, R_s is the sheet resistance and k is the geometric correction factor. A correction factor of 3.5749, attained from a table of solved equations for various geometries, was used to account for the rectangle dimensions and geometry.[88]

Typically, the e-beam evaporated films had a higher measured resistivity (0.2145 Ω μm) than sputtered films (0.173-0.2066 Ω μm) and of bulk nickel resistivity (0.0699 Ω μm). The resistivity of a sputtered thin film can range from equal to bulk value to an order of magnitude larger than bulk (e.g. Misra *et al.*[89]), though no reference to thin film nickel resistivity was found. Nevertheless, this was somewhat expected since the thin-film nickel has much more granularity than bulk material and at the base pressures and deposition rates used ($1 \cdot 10^{-6}$ and above for e-beam evaporated films and $3 \cdot 10^{-7}$torr for sputtered films, and a few Å/s) and a sizable portion of the highly reactive nickel may have oxidized during deposition. This is consistent with the fact that, generally, the e-beam evaporator was pumped to a higher base pressure than the sputtering system (simply due to its diminished pumping ability due to a leak/oil backstreaming).

5.4.3 Photolithography

Lift-off was chosen as the method of patterning mainly due to the possibility it brought of sharp edges of e-beam evaporated materials and multiple layers deposited and patterned accurately with a single photoresist pattern (as opposed to chemical etching that is selective for each material). Also, with lift-off, there is a lower probability of contamination of the junction area when compared to chemical methods or sputter etching means of patterning. Chemical dry etching was eliminated as an option as Ni is not chemically etched by any of the plasma gases available at the University of Alberta Nanofab (and it is not known if *any* plasma chemically etches Ni).

Contact photolithography was used with a bi-layer HPR 504 photoresist and LOR 5B lift-off resist. The procedure used was, in general, standard to the photoresists used. Nevertheless, the 1μm resolution of the 'small features' mask pattern was never achieved in practice. Typically 3-4μm lines (channels) were developed from exposure to 2μm line mask patterns, but the 1μm mask patterns often produced incompletely developed channels (Figure 22). Better resolution may have been possible with a thinner photoresist thickness (1.2μm thick HPR-504 and 400nm LOR resists were used) but the spin speed was already fairly high (4000rpm for the HPR-504) and an alternate thin resist was not readily available. The process parameters of 4s exposure and 20s development were found to be optimal for the photoresists used and no improvements in feature definition were seen either by increasing or decreasing the exposure or development times (2s/6s exposures, 10s/40s/60s developments). Rather, features were significantly degraded with shorter exposure and development times. Even with these parameters it can be seen (Figure 22) that the smallest (1μm) features were not well defined (underdeveloped or underexposed), while at the same time the larger features had rounded corners and were larger than their nominal sizes (overdeveloped or overexposed).

Figure 22 Image of an MIM diode with a patterned 5μm (2μm mask pattern) second electrode layer of photoresist as seen through an optical microscope (left) and chronic pattern underexposure for a 5μm (1μm mask pattern) line (right).

The procedure for standard photolithography at the University of Alberta Nanofab using HPR 504 was followed. The UV lamp of the mask aligner was allowed to warm up for at least 30 minutes prior to each use and an exposure time of 4s was used for the 'Elmo' or 'Oscar' mask aligners. Pre-bake of the HPR-504 was done for 60s on a vacuum hot-plate at 110°C. Despite varying the lithography parameters, as previously mentioned, the resolution of the mask feature sizes was not achieved. Nevertheless, it was deemed not necessary to optimize the lithography any further since a minimal feature size lithography with sharp corners is not a requirement for successful fabrication of the MIM devices. As a result of larger (than mask) features it is expected that the capacitance is higher and therefore the devices would have a lower corner or maximum frequency of operation.

Pre[illegible]ons to minimize particulates on the photomask were taken with cleaning of the mask with acetone and DI water (approximately 2 minute rinse each) before each use and regular soaks in cold piranha solution for 5 minutes, followed by a 5 minute rinse in DI water. Precautions to minimize particulates on the photoresist-covered substrate included covering the photoresist spinner with a metal foil covered plate (which minimized particulate attraction by electrostatic force created during spinning) and blowing the substrate and mask with a nitrogen gun before exposure.

For 'clean' removal of metal in the lift-off process it was necessary to have an overhanging structure such that there is no physical link between metal deposited on the wafer surface and the metal deposited on the photoresist surface. This type of overhang structure is commonly done using chemical treatment of the photoresist surface prior to development with chlorobenzene or toluene, as in the process developed by Hatzakis *et ---al.*[90] The process involves soaking the photoresist covered wafer before or after exposure to the UV through a photomask. While the chlorobenzene photoresist treatment was attempted it resulted in a large decrease in resolution in the photoresist used in the Nanofab for micrometer size features. Several soak times were attempted ranging from 30s to 6 minutes. None of the attempts produced features that were not overly rounded or poorly defined. Another lift-off method is to use two layers of different materials (photoresists) with different etch characteristics. This method proved much simpler and successful. It involved using a layer of resist (Microchem LOR 5B, Nanofab supplied) that is etched by a different base (TMAH) than that which develops HPR-504 (KOH). After development of the HPR-504 top layer and a two-minute hard-bake the sample is dipped in the TMAH based developer (MF-321) for 45-60s etching away the LOR layer. Images of the overhanging structures are shown in Figure 23 and Figure 24. Figure 23 displays a 3μm overhang of the HPR-504/LOR 5B bi-layer patterning the bottom electrode, as viewed by a CCD camera on an optical microscope.

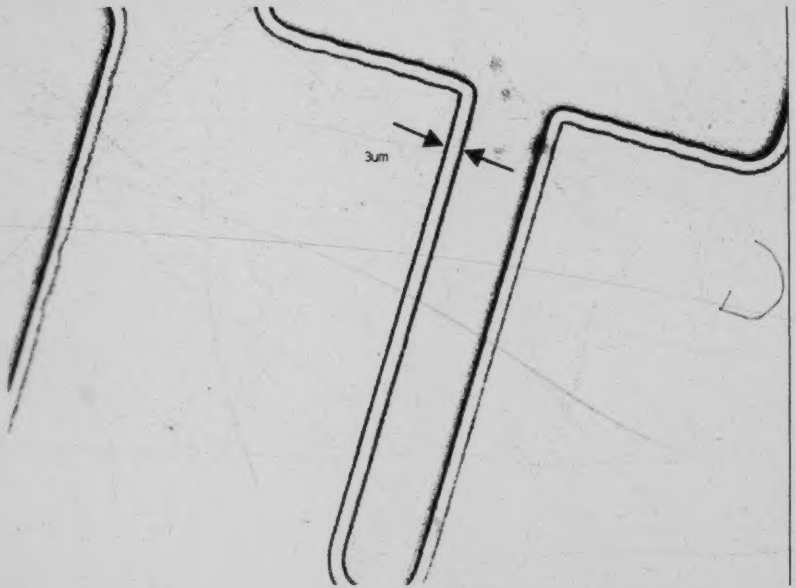


Figure 23 Overhead view of a 3µm undercut. Outer pattern is LOR, inner pattern is photoresist on top of the LOR.

Figure 24 displays an SEM image of a smaller overhang (0.5µm). The LOR and HPR 504 thicknesses are the measured values (443.3nm and 1.253µm, respectively).

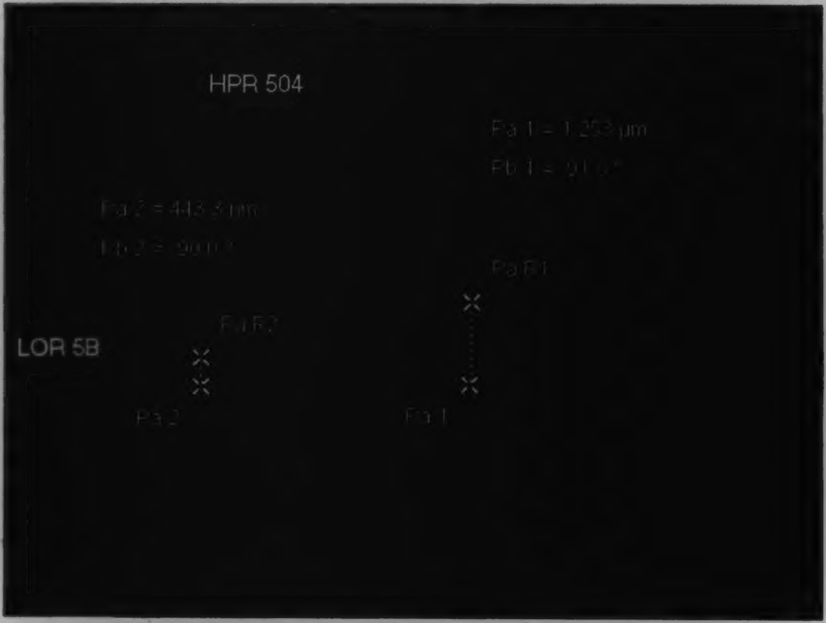

HPR 504
Pa 1 = 1.253 µm
Pb 1 = 91.0 °
Pa 2 = 443.3 nm
Pb 2 = 90.0 °
Pa R1
Pa R2
LOR 5B
Pa 2
Pa 1

Figure 24 Isometric view of overhanging photoresist

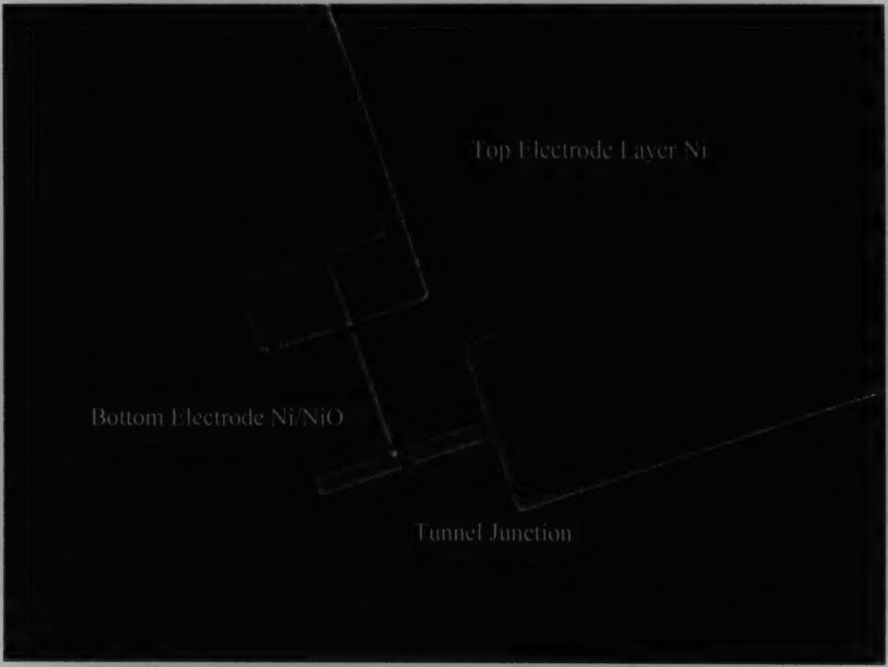


Figure 25 Optical image of completed MIM diode

5.4.4 Ellipsometry

Overview of Ellipsometry

Ellipsometry is a non-destructive method of measuring several optical and physical properties of a bulk or thin film sample using polarized light. Using appropriate mathematical models ellipsometric measurements can determine the optical constants n and k (refractive index and extinction coefficient), surface roughness, thin film layer thickness, stoichiometric composition and optical anisotropy. Ellipsometry is particularly attractive for thin film thickness measurement since it is precise to Ångström resolution. Ellipsometric measurements can be made at a single wavelength with a laser ellipsometer or a range of wavelengths with a spectroscopic ellipsometer. The basic concept of ellipsometry is shown in Figure 26.

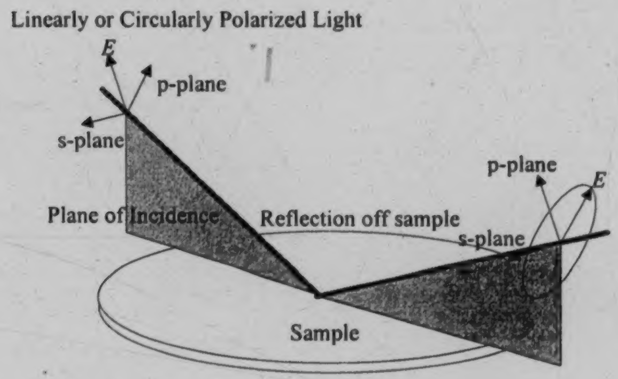

Linearly or Circularly Polarized Light
E
p-plane
s-plane
Plane of Incidence
Reflection off sample
p-plane
E
s-plane
Sample

Figure 26 Elliptically polarized light from reflection off a sample

A linearly polarized beam of light (*s* and *p* components of the beam are in phase or 180° out of phase) is emitted from a collimated light source and is incident on a sample to be measured at an oblique angle. The reflection causes a change in relative phase and amplitude between the *s* and *p* components and so the wave becomes elliptically polarized. The change in relative amplitudes and phase is given by $tan(\psi)$ and Δ, respectively. The complex reflectance coefficient ρ is measured, related to the parameters ψ and Δ and reflectance coefficients in the s and p planes R_p and R_s by:

$$\rho = \tan(\psi)\exp(i\Delta) = \frac{R_p}{R_s} \tag{33}$$

The relative amplitudes of the component *s* and *p* polarized waves are determined by measuring the 'height' of the signal divided by its 'width', defined by the largest dimension and smallest dimension of the ellipse, respectively, as measured from the centre. Figure 27 shows a typical laser ellipsometer set-up. There are several configurations of ellipsometers but both the laser and spectroscopic ellipsometers used in this research were the rotating analyzer type. The laser (or lamp in the spectroscopic case) provides a beam of collimated light, which is then linearly polarized by the polarizer. In the spectroscopic case an additional component, the monochromator, filters out all wavelengths except those of interest at the output of the light source. The light beam then passes through a compensator, which changes the relative phase between the *s* and *p* component polarizations. If the compensator exists in the particular ellipsometer setup then it is a quarter wave or 90° phase shifter resulting in a circularly polarized beam. The beam then reflects off the sample surface in an elliptically polarized state. The reflected beam then passes through another polarizer (known as an analyzer for this function), which rotates, thereby selectively passing the beam at different angles. The detector, a light intensity-metering device, simply measures the light intensity that is passed through the analyzer. The light induced current is given by:

$$I_D = I_0(\cos A \cos P \rho_d + \sin A \sin P \rho_v) \qquad (34)$$

where I_0 is the average light intensity, A is the analyzer angle, P is the polarizer angle and ρ_d is the complex ellipsometric reflectance ratio, to be solved for, and ρ_v is the complex compensator transmittance. The ellipsometric parameters are extracted from the above equation by Fourier expansion.[97]

Once the ellipsometric parameters ψ and Δ are determined they are used to fit to a chosen material and physical property model or combination of models. The three-phase model (ambient-thin film-substrate) model was used to fit the parameters with the data.

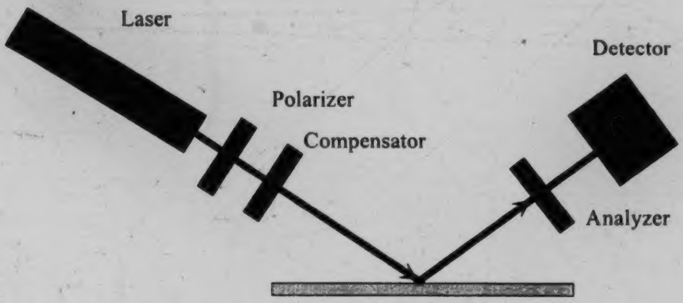

Laser
Detector
Polarizer
Compensator
Analyzer

Figure 27 Diagram of a typical laser ellipsometer

Three-Phase Model

The three-phase model describes the most common configuration used in ellipsometric measurements. Typically, a thin film sits on a substrate and the reflected light is measured in ambient air.

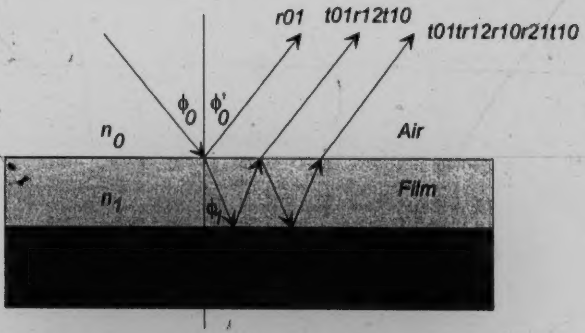

r01
t01r12t10
t01tr12r10r21t10
ϕ0
ϕ'0
n0
Air
n1
Film
ϕ1

Figure 28 Reflection of light at thin-film on substrate

Figure 28 shows the light interaction relationships in the three-phase models. The Fresnel reflection coefficients (r_{xy}) are given by:[91]

$$r_{xy}^{p} = \frac{n_y \cos(\phi_x) - n_x \cos(\phi_y)}{n_y \cos(\phi_x) + n_x \cos(\phi_y)} \tag{35}$$

$$r_{xy}^{s} = \frac{n_x \cos(\phi_x) - n_y \cos(\phi_y)}{n_x \cos(\phi_x) + n_y \cos(\phi_y)} \tag{36}$$

The incident beam is reflected and transmitted through the film many times and the total reflected beam is a sum of the individual paths re-emerging from the surface. If the individual paths are products of their interactions (i.e. path 1 is simply a reflection off the surface r_{01}, path 2 is two transmissions through the thin film and a reflection off the film/substrate interface) then the total beam is:

$$r_{tot} = r_{01} + t_{01}r_{12}t_{10}e^{-i2\beta} + t_{01}r_{12}r_{10}r_{12}t_{10}e^{-i4\beta} + \ldots \quad (37)$$

Which converges to:

$$R^p = \frac{r_{01}^p + r_{12}^p e^{-i2\beta}}{1 + r_{01}^p r_{12}^p e^{-i2\beta}} \quad (38)$$

$$R^s = \frac{r_{01}^s + r_{12}^s e^{-i2\beta}}{1 + r_{01}^s r_{12}^s e^{-i2\beta}} \quad (39)$$

$$\beta = 2\pi\left(\frac{d}{\lambda}\right)n_1 \cos\phi_1 \quad (40)$$

for s and p polarizations.[92] As can be seen from the above equations R^s and R^p are dependent on *d*, the thin film thickness, so *d* can be indirectly measured with ψ and Δ measurements.

The relationships between angles ϕ_0 and ϕ_1, and ϕ_1 and ϕ_2 are defined by Snell's law:

$$n_0 \sin(\phi_0) = n_1 \sin(\phi_1) = n_2 \sin(\phi_2) \quad (41)$$

while incident and reflected angles are equal. Typically, using the above equations the film thickness is initially guessed and then it, and the film's index of refraction, are adjusted such that the resultant ψ and Δ agree with the measured data (see next section).

Ellipsometric Characterization of Films

A critical part of each process was the reliable and reproducible growth of oxide on the metal (nickel) to specific thickness. The thickness of the oxide covering the metal was measured with a laser source rotating analyzer ellipsometer (Gaertner Model L125B λ=632.8nm[93]) that was checked against a variable angle spectroscopic ellipsometer (J.A. Woollam VASE[94]) by measurement of a sample of Ni/NiO at an incident angle of 70° from the normal to the sample and comparing ψ and Δ values.

Values for the optical constants n and k were assumed (at λ=632.8nm Ni: n=2.05-2.3 k=-3.8[95,96], NiO: n=1.95 $k\approx 0$[97,98]) since it was not possible to determine their values in-situ immediately after deposition and once removed from vacuum the metals are covered with a layer of native oxide. The assumed values were compiled from data presented by the given references and the values for Ni n,k = 2.05, -3.8 and NiO n= 1.95 were used in each measurement.

Unfortunately, with ellipsometry the thickness of a film can not always be determined analytically from measured ψ and Δ values if the n and k values are not explicitly known. Fitting techniques must be employed to extract values for optical constants n, k and film thickness. Generally, this fitting is done automatically by the ellipsometry control and analysis software on a PC connected to the ellipsometer itself. Manual fitting can be done by minimizing the (Root) Mean Squared Error (MSE), given by:

$$(R)MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{\psi_i^{\mathrm{mod}} - \psi_i^{\mathrm{exp}}}{\sigma_{\psi,i}^{\mathrm{exp}}}\right)^2 + \left(\frac{\Delta_i^{\mathrm{mod}} - \Delta_i^{\mathrm{exp}}}{\sigma_{\Delta,i}^{\mathrm{exp}}}\right)^2\right]} \qquad (42)$$

where *mod* and *exp* denote model and experimental values, respectively, and σ_{ψ} and σ_{Δ} are the respective standard deviations of ψ and Δ.

The software used for measurement in this thesis were Gaertner's *2WS* and *GEN* programs. While the minimization function of these programs is not explicitly known it is likely that the MSE function, above, is used in these programs, since comparison with the 3 layer model yielded the same film thicknesses when manually calculating the function and comparing with the 2WS and GEN programs. For instance, a curve for the psi-delta relationship for NiO on Ni as a function of oxide thickness at an incident angle of 70° is given in Figure 29 (for values of $n_{Ni} = 2.05$, $k_{Ni} = -3.8$, $n_{NiO}=1.95$). When a given psi-delta measurement is taken and the psi deviates from the graph by a greater amount than the delta it can be assumed that the minimization will be dominated by minimizing the psi error. Using this method, the minimal error point on the graph is found and used to determine the film thickness.

Reproducibility of the ψ, Δ measurements was, in general, better than ±0.25° for ψ and ±0.5° for Δ, with most measurements having reproducibilities of just ±0.1° for ψ and ±0.25° for Δ, or better. In the worst case this represents approximately ±3Å in film thickness. These measurements were determined by comparison of repeated measurements at single measurement sites.

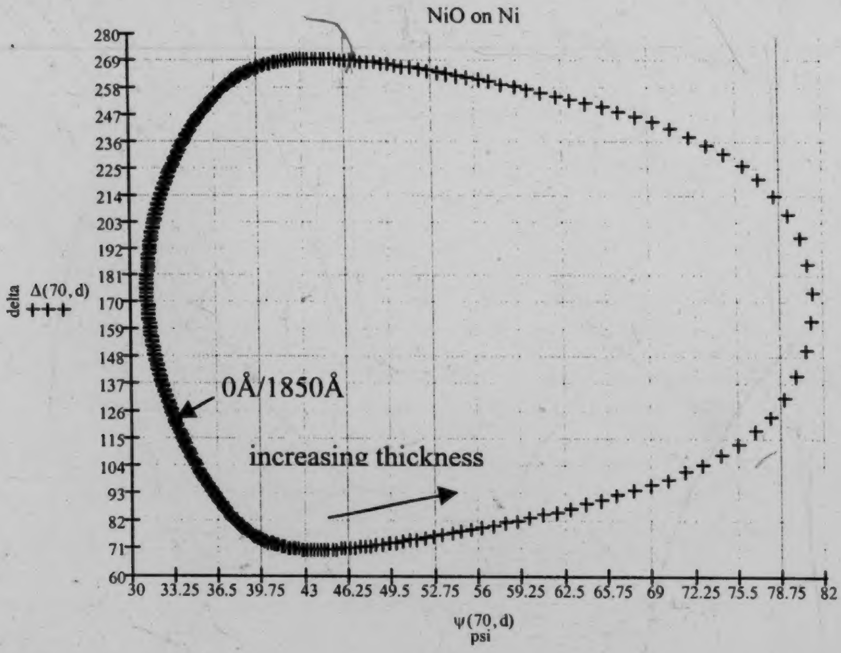


Figure 29 Psi-Delta curve for NiO on Ni from 0 to 1850Å, as calculated from equations 33,38-41.

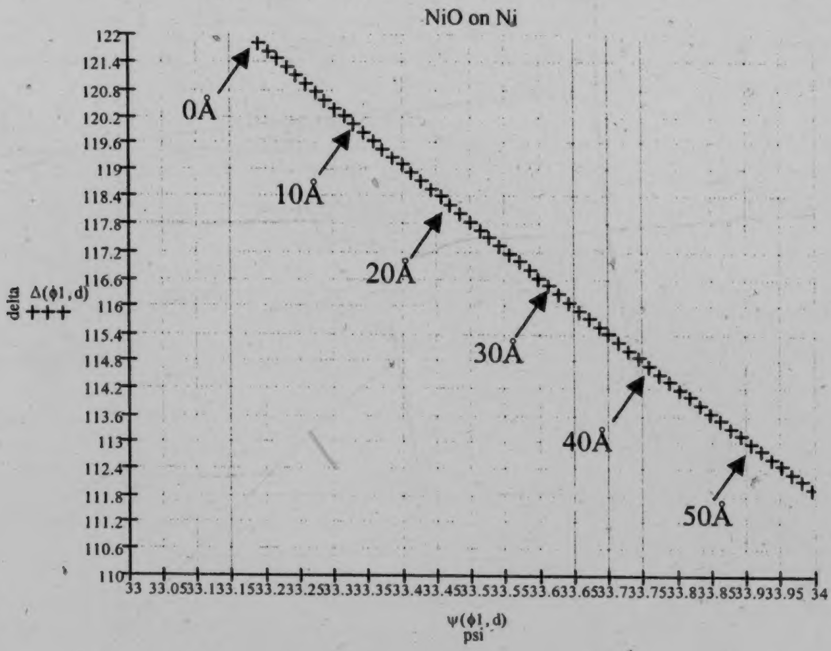


Figure 30 Psi-Delta curve (zoomed) to 0 to 56Å (each cross = 1Å)

Oxide Removal and Growth

Oxide growth is a critical process in the development of an MIM diode. It requires oxides with thicknesses of tens of Angstroms to be grown reliably from sample to sample. With this in mind it is necessary to first remove any contaminants on the surface of the metal and the native oxide present.[26] Also, the removal of the native oxide creates a common reference for each oxidation run and ensures that the final oxide is due to the plasma oxidation, rather than a combination of native oxide growth in ambient air and plasma oxidation. The use of an Ar+ sputter-etch pre-clean step was found to greatly increase the reproducibility of niobium tunnel junctions by Broom *et al.*[75] Further, the removal of the native (or thermal) oxide initially present on Ni films ensures that the grown NiO is characterized by the plasma oxidation process and not by a combination of thermal and plasma oxidation (which could be difficult to control). Native oxide has been found to inhibit plasma oxide growth on Pb electrodes.[99] To prevent a similar condition (without knowing *a-priori* whether this condition is true for Ni) the oxide is first removed, then an oxide is grown without breaking vacuum.

Several chemicals (BOE, acetic acid, acetone, nitric acid, ammonium hydroxide, KOH and based TMAH developers) were tested for their effects on the NiO. The NiO thickness on the test sample was measured by ellipsometry before and after soaking (and rinsing and drying at a wet bench for approx. 1 min each) the sample in the test chemical. The purpose was not only to find a chemical to remove the native oxide but also to ensure that none of the chemicals used in other steps of the process unintentionally strips or grows the oxide. Of the chemicals tried only ammonium hydroxide had any significant effect on the oxide thickness. Ammonium hydroxide at room temperature removed a portion (~50%) of the thickness of the native oxide of a sample of nickel initially covered with 30Å of native oxide. A more effective method was to expose the sample to medium-high power intensity Ar^+ plasma (~100W) for 10-20 minutes. The native oxide on the nickel surface is sputtered by the Ar^+ plasma, ultimately leaving the metal with a clean unoxidized surface. This technique was characterized with ellipsometric measurements (Figure 31) taken at 5-minute intervals of plasma exposure. After each interval the sample was removed from the chamber and measured with a laser ellipsometer at two points; an edge and the centre of the unpatterned sputter deposited nickel wafer. The oxide sputter removal process is approximately linear with time until it is limited by the small quantity of oxide left on the sample (equilibrium at a few Å). The different thicknesses measured (particularly at the final thickness) for the two measurement sites suggest a variation of optical constants between the two points for reasons that are not obvious. Nevertheless, the final thickness of the removed oxide converges to within a few Å, suggesting that the assumed values of the optical constants are sufficiently close to the actual values that meaningful observations can be made. While the ultimate thickness of the oxide may not have been 0Å, as shown in the graph, it may be impossible to measure 0Å thickness if the sample is exposed to air. NiO forms a 1nm thick layer on Ni when exposed to air within a matter of seconds. However, if the sample is not removed from the chamber it can be assumed that the NiO thickness is sufficiently close to 0Å that thermal oxide will not play a significant role in subsequent plasma oxide formation.

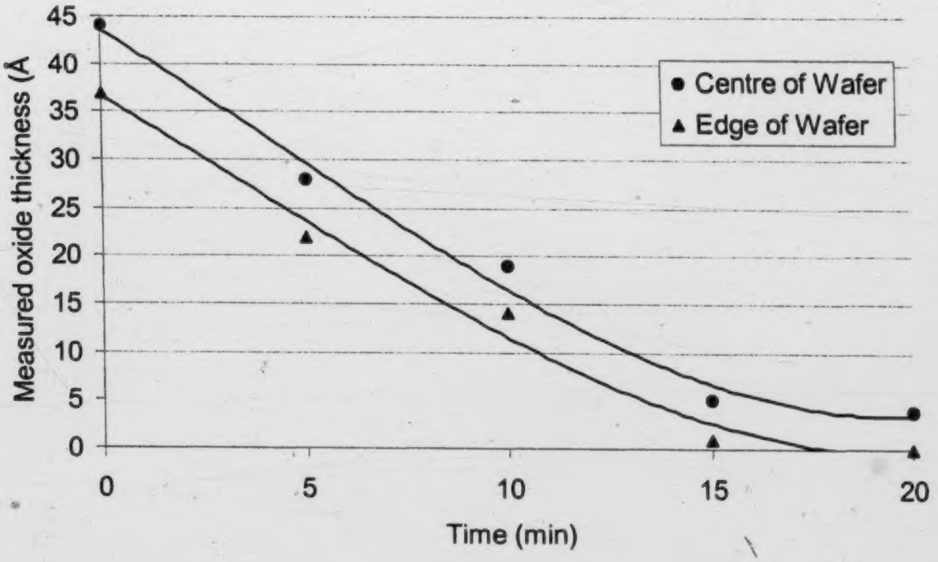


Figure 31 NiO removal from sputter deposited Ni by Ar^+ plasma (107W) exposure in an RIE chamber

Once the sample is cleaned of native oxide a metal oxide layer is grown. This can be done with exposure to ambient air, oxidation by anodization or oxidation in oxygen plasma. The oxide growth is expected to be a diffusion process driven by a contact potential between the metal and the adsorbed oxygen, as described by the Cabrera-Mott model for oxide growth.[100] If ambient oxidation is used, the oxide is usually grown such that the diffusion reaches a steady state to ensure minimal amounts of pinholes. When using this method the choice of oxide thickness is limited. If an oxygen plasma is used the oxide thickness can be prescribed by interrupting the oxidation before it reaches steady state. For reproducibility a time independent process was desired. One such process was proposed by Greiner for the fabrication of Josephson junctions.[101] The process, in which the metal surface is simultaneously oxidized by the oxygen plasma and sputter-etched by the argon plasma, was applied to the RIE chamber at the University of

Alberta Nanofab and was found to give repeatable results of oxide thickness as a function of both oxygen partial pressure and R.F. power (i.e. using the same parameters repeatably yielded oxides of the same thickness as measured by ellipsometry). The oxide thickness was found to be stable after approximately 5 or 6 minutes (changing less than a few percent between ellipsometric measurements of samples after exposure to the $Ar^+:O_2$ plasma exceeded 5-6 minutes) confirming that the process is time independent. Figure 32 shows the relationship between O_2 pressure, R.F. power and the resulting oxide thickness on Ni film (as measured by ellipsometry). Clearly, oxide thickness increases as O_2 pressure or R.F. power is increased. This is expected since, as the oxygen pressure is increased, the relative amount of oxygen is increased and the sputtering gas (Ar) is decreased, and subsequently the amount of oxidation increases relative to the amount of sputtering. The power dependence of oxide thickness is not as obvious. In fact, it was found by Greiner[101] that an increase in power would decrease the resistance of the junctions, suggesting more pinholes presumably due to a thinner oxide. Greiner explained this by suggesting that the sputtering is more strongly power dependent than the diffusion based oxidation. The results here show the opposite effect wherein the oxidation increases at a faster rate than sputtering with increased R.F. power levels. However, the power dependence of the oxide barrier thickness is a weaker dependence than the oxygen partial pressure, as seen by the smaller difference in oxide thickness between higher barrier pressures. Nevertheless, the process is a 'prescribed' process as the oxide barrier thicknesses are repeatably grown to the thicknesses shown in Figure 32, between samples.

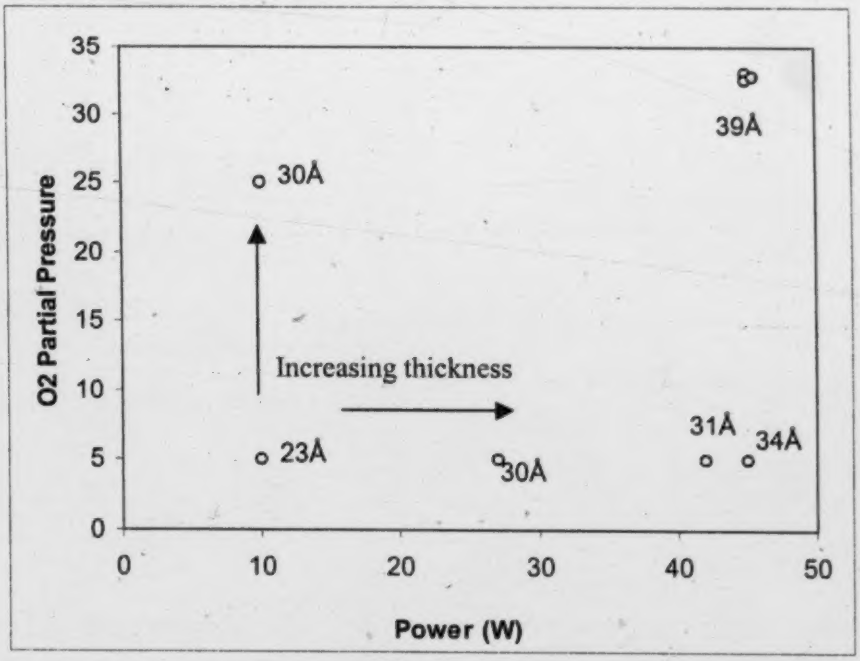


Figure 32 Pressure and power dependence of the Ar^+:O_2 sputter-etch oxidation technique

6

Chapter 6: D.C Results and Analysis

6.1 Chapter Overview

In this chapter the electrical characteristics of the fabricated devices are presented. The measured data are interpreted in terms of Simmons' theory for tunnel junctions. From the theoretical fits such parameters as the work-function difference ϕ_B, effective junction area *JA*, and the effective insulator thickness *d* can be extracted. Tunnelling is shown to be the dominant current mechanism by characteristic F-N plots of the devices in addition to the non-linear *i-v* curves.

6.2 DC Testing of MIM diodes

The microfabricated MIM diodes were tested at D.C. with an HP 4145A semiconductor parameter analyzer (SPA) connected to the devices via a Wentworth Pro 0195LH probing station (Figure 33).

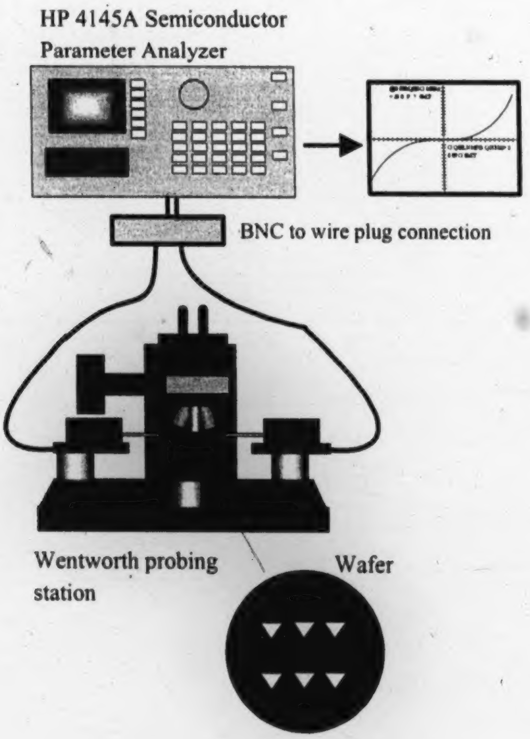

HP 4145A Semiconductor Parameter Analyzer
BNC to wire plug connection
Wentworth probing station
Wafer

Figure 33 D.C. Characterization set-up

A minimum voltage sweep of –1V to 1V was applied to the devices to a maximum current of 10-50mA and the sweep was increased until a sufficient non-linearity was observed or the device failed (device failure determined by a sharp decrease in resistivity and linear *i-v* characteristics, see Figure 34). Since the devices were based on very thin oxides there was a threat of electrostatic discharge when connecting the devices to measuring equipment. Precautions were taken by connecting a high-resistivity (>100kΩ) shunting resistor between the two probe leads prior to contacting the probes to the device pads.

The first functional devices were achieved with run 1. The resultant diodes were weakly non-linear and did not survive overnight storage before more thorough measurements could be made. An example of a diode from this run, measured directly after top layer lift-off and again on the following day (shorted diode trace), is given in Figure 34.

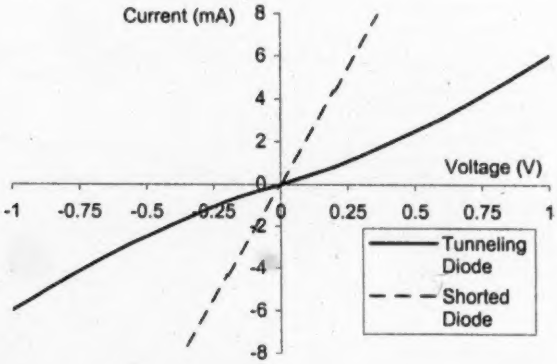


Figure 34 Tunnel *i-v* diode characteristics from run 1 (~30Å oxide thickness) before and after dielectric breakdown

A functional device yield of almost 40% was achieved on run 3. However, the devices had varying characteristics, most likely due to process non-uniformity and statistical variation in oxide thickness (variation due to random impurities, surface roughness, etc.). Also, the diodes had unexpectedly high transition voltages between linear and non-linear characteristics. No obvious pattern was apparent in the placement of diodes on the wafer and diode performance.

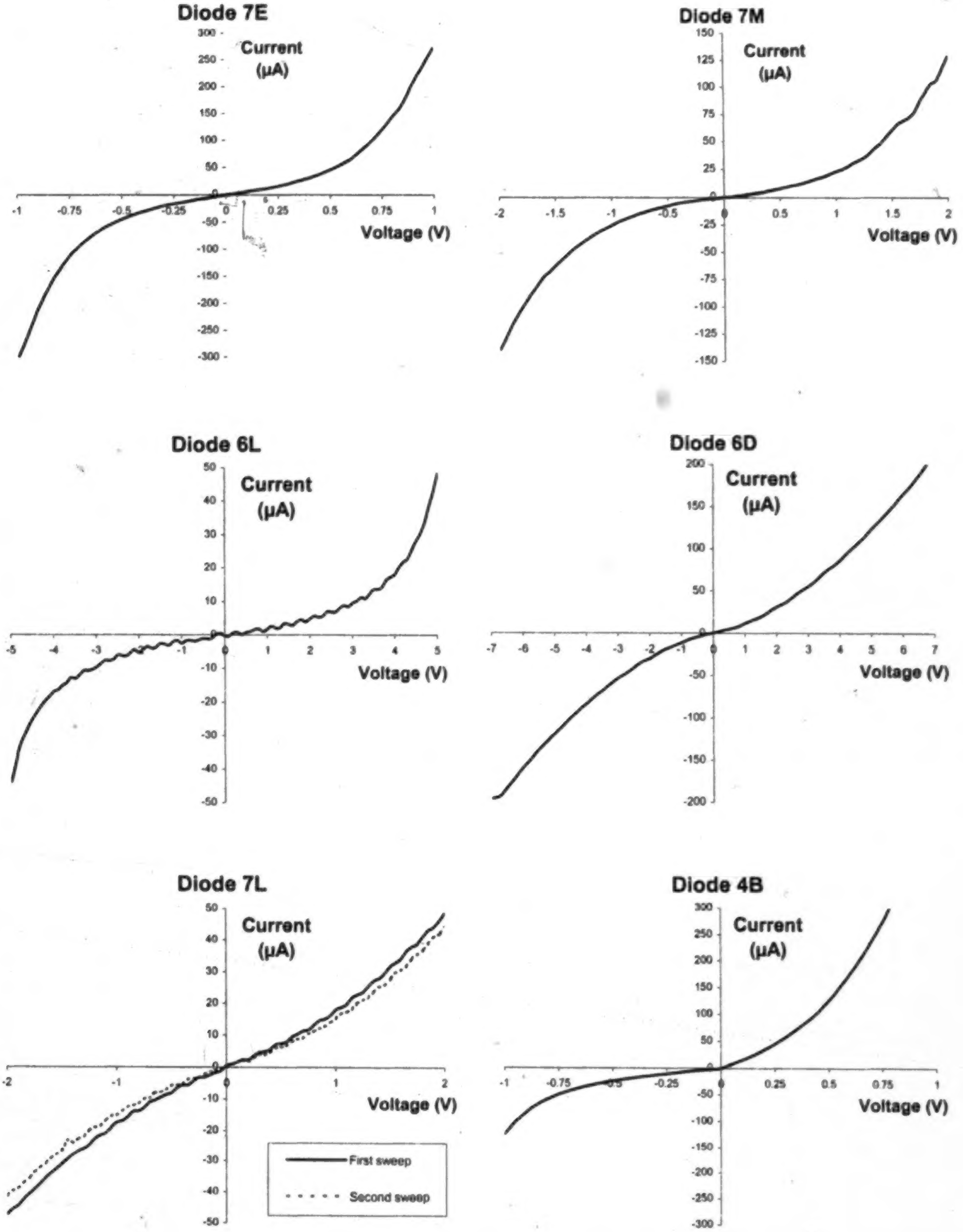
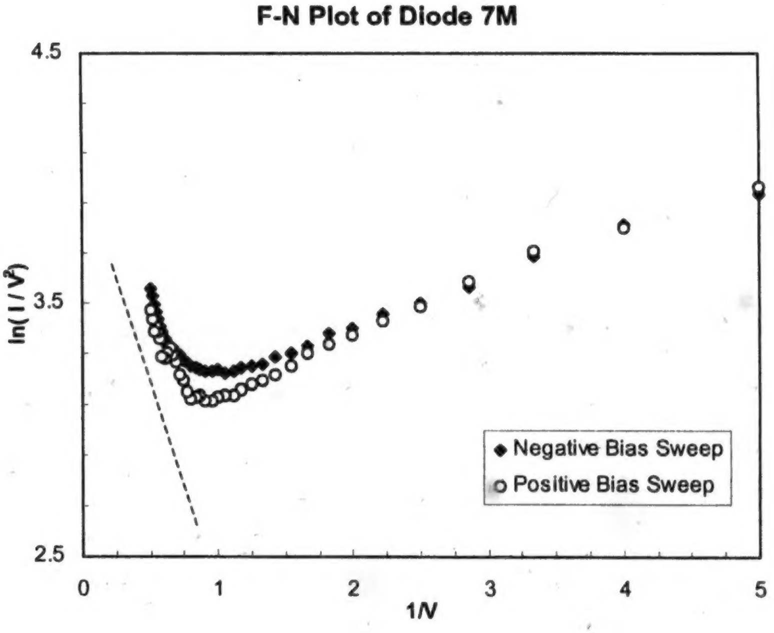


Figure 35 ***i-v*** **characteristics of 40Å (nominal) barrier diodes from run 3**



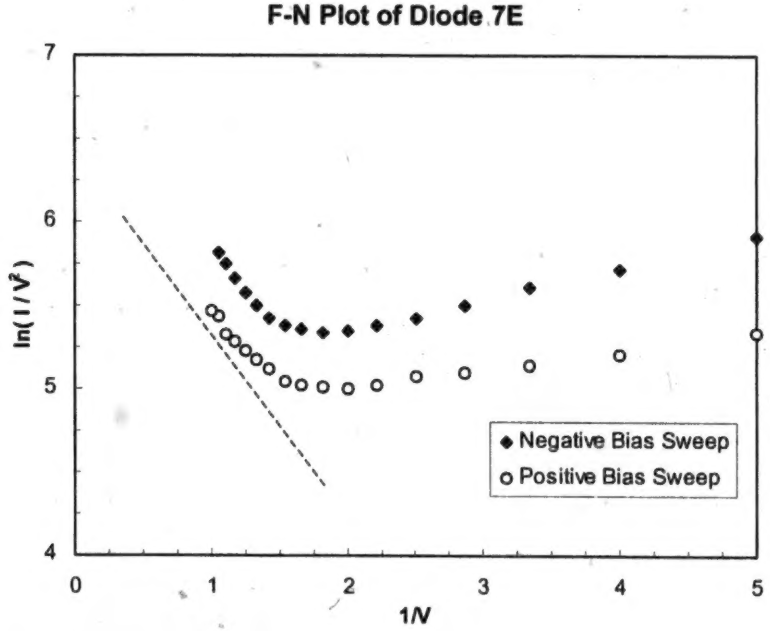


Figure 36 Fowler-Nordheim plots of selected diodes. The diodes asymptotically approach the F-N regime (linear, negative slope F-N plot characteristics)

Table 1 Resistances of measured diodes

Device:	4B	4C	6D	6L	7E	7L	7M
Zero Bias R (ohms)	672	25.66k	98.93k	362.1k	17.7k	48.1k	110.9k
Approx. Dynamic (Video) Resistance	309	8.0k	35.6k	114.5k	3.16k	41.8k	14.3k

Typical measured *i-v* characteristics for 40Å (nominal) Ni-NiO-Ni devices from run 3 are shown in Figure 35. A wide range of voltage ranges was measured for the diodes before non-linearity was observed (from 1V to 7V). Stability of the diodes also varied with a few devices changing *i-v* characteristics (non-linearity, current magnitude) between applied voltage sweeps. Diodes 7L and 4B noticeably changed characteristics between voltage sweeps (7L) or within a voltage sweep (4B). These diodes showed less non-linear curves after the transition and either stabilized (7L) or deteriorated further (4B).

Corresponding Fowler-Nordheim (F-N) plots are shown in Figure 36 for two of the diodes. The Fowler-Nordheim plot is typically a plot of $\ln\left(\frac{I}{E^2}\right) \text{vs} \frac{1}{E}$ but since the true tunnel barrier thickness was not known plots of $\ln\left(\frac{I}{V^2}\right) \text{vs} \frac{1}{V}$ are given. Through examination of the Fowler-Nordheim region relation (Equation 1) it can be seen that the F-N plot should yield a negatively sloped line for this region. Through comparison of the

F-N plots (which asymptotically approached a linear trace for high bias) it is determined that the diodes approached the F-N regime but the current was not dominated by Fowler-Nordheim tunnelling phenomena. This is particularly true of the less non-linear diodes (i.e. run 1 diodes, run 3 diodes 7L, 6D), which were not measured at enough of a high bias so that an asymptotic trace was viewed and hence no F-N tunnelling was evident. A similar condition was observed by Scherer *et al.*[31] and they were able to verify the work function from the graph and build SET devices constructed from similar tunnel junctions. The slight shifts in data traces between the negative and positive bias sweeps indicates a slight asymmetry in i-v characteristics, likely due to the difference in interfaces between the bottom electrode and the oxide grown on its surface and the sharper transition between the top electrode and the oxide, on which it is deposited.

Nevertheless, the shapes of the F-N plots corroborate the theory that tunnelling is the cause of the non-linear *i-v* characteristics.

6.3 Yield

The device yield varied among the various process runs from zero (all shorts or opens) to over 1/3. In total, five runs were run to completion of devices, three runs of which did not produce any functional devices. The first two runs that did not produce devices suffered from problems with the lift-off lithography (the first due to internal stress in the metal, which may have been caused by opening the chamber too soon to relatively cool ambient air; the second due to poor e-beam film adhesion, at least partially caused by poor e-beam system performance (oil backstreaming – discussed later)). The problem was made worse by the poor adhesion of the nickel to the aluminum adhesion layer. Film adhesion problems were reported by at least one other researcher using the e-beam evaporator depositing titanium on silicon (a combination which should be very adhesive).[102] The consensus amongst users was that the e-beam pumping system was to blame as evidence of oil backstreaming was found on the gate valve. The system has since been refitted with a cryogenic pump and no recent adhesion problems have been reported by Nanofab staff.

Film thicknesses were measured with a mechanical profilometer (Alpha Step 200) for both sputtered and e-beam evaporated films. The e-beam evaporator was also equipped with a crystal thickness monitor (using 6MHz Au crystal resonators from Phelps Electronics Inc, P/N:102-03) which was checked against values from the mechanical profilometer and was found to be a fairly accurate guide (within 10% of profilometer values) for *in-situ* deposition thickness monitoring. The deposition rate of nickel by sputtering was determined to be 3.3Å/s (by profilometer measurement of a 10 minute / 300W nickel sputter deposition which yielded a 200nm thick film). All of the runs listed here had sputtered Ni bottom electrodes at 300W. Only the top electrode deposition methods and parameters had changed from run to run (see Table 2 and Table 3)

Table 2 Summary of all sputtering run parameters

Run	Characteristics	Bottom Electrode Deposition Base Pressure (torr)	Top Electrode Base Pressure (torr)	Top Electrode Shutter Open (min)	Top Electrode R.F. Power (Open Shutter) (W)
1	Some functional devices lasting less than 24h	$4.8 \cdot 10^{-7}$	$3.2 \cdot 10^{-7}$	1 for Al, 10 for Ni	300
2	Top layer lifted off	$4.8 \cdot 10^{-7}$	$2.4 \cdot 10^{-7}$	2/1 for Al, 2/10 for Ni	50W/200W for Al, 200W/300W for Ni

Table 3 Summary of e-beam run parameters (deposition rates are for top electrode)

Run	Characteristics	Bottom Electrode Deposition Base Pressure (by sputtering) (Torr)	Base Pressure before any Ti gettering (Torr)	Electrode Deposition Base Pressure Al (Torr)	Electrode Deposition Base Pressure Ni (Torr)	Deposition Rate Al (Å/s)	Deposition Rate Ni (Å/s)
3	Many Functional Devices	$1.3 \cdot 10^{-7}$	$1.8 \cdot 10^{-6}$	$7 \cdot 10^{-7}$	$2 \cdot 10^{-6}$	5	1.5
4	No Devices/Metal detached	$5 \cdot 10^{-7}$	$4.0 \cdot 10^{-6}$	$1.4 \cdot 10^{-6}$	$2.8 \cdot 10^{-6}$	8	1.5
5	All shorts	$5 \cdot 10^{-7}$	$4.3 \cdot 10^{-6}$	$8 \cdot 10^{-7}$	$1.4 \cdot 10^{-6}$	4	2.2

6.3.1 Run 1

In the first run, sputtering with uniform power of 300W was used to deposit the second electrode. A thin (approx 10nm) Al layer was first deposited, then a thicker (200nm) layer of Ni was deposited on top of the Al in the channels on the lift-off patterned wafer. A thickness of 30Å of NiO was grown for this run in the RIE. For lift-off of the second electrode only a rinse with a low-pressure DI water stream was used. This run produced functional devices (albeit with plenty of debris) that had slight non-linearity and lasted less than 24 hours. An attempt to clean the substrate, by completing the lift-off (by immersing the wafer in acetone and rinsing in the dump-rinse for 5 minutes) resulted in the top features washing away before more measurements and characterization could be performed.

6.3.2 Run 2

Since run 1 had produced some functional devices it was theorized that improved characteristics could be attained by changing the deposition parameters of the second electrode. The devices of run 1 were only slightly non-linear and a higher degree of non-linearity and electrical stability was desired. It was thought that pinholes and hot-spots had made the diodes of run 1 unstable. A better quality oxide (fewer pinholes, more uniform interface) could be achieved with different deposition parameters. Specifically, a lower power was used such that the incoming flux of deposited material was less energetic. The oxidation parameters were kept constant from the previous run. A thickness of 30Å of NiO was grown in the RIE at an oxygen partial pressure of 5% at a total pressure of 0.1 torr and a plasma RF power of 42W (45W input / 3W reflected). To avoid damage of the thin oxide the second electrode was initially (first 5 min) deposited with a lower power before switching to a higher power (for faster deposition). This strategy failed, in that features did not stick to the substrate (or the photoresist) as the metal 'curled' and detached after deposition. This was likely due to internal stresses in the metal introduced by varying the R.F. power during the deposition, thereby effectively producing two layers of the same metal with different characteristics (grain size, density) or difference in metal composition between the two metal depositions (e.g. increased adsorbed oxygen content at lower deposition rate).

6.3.3 Run 3

In the following run, a change in several parameters led to the most successful wafer of devices. An oxide thickness of 40Å, measured by ellipsometry, was grown in the RIE using a 33% oxygen partial pressure at a total pressure of 0.1 torr and 42W. The 10Å increase in oxide thickness for this run was included to minimize the probability of pinhole shorts. The sample was also left for 4 days in ambient air prior to second electrode deposition, allowing thin pin-hole sites to be oxidized naturally. The average oxide thickness did not change over the 4 days, as measured by ellipsometry.

E-beam evaporation was used to deposit the second electrode, partly due to the sputtering system becoming unavailable (for several months for maintenance investigation into the cause of the very poor adhesion) but also due to the fact that e-beam evaporation is, in general, a more gentle deposition method. E-beam evaporant flux has an energy that is in the 0.1eV range, while sputtered material flux has an energy in the

10-100eV range.[103] This would also explain the poorer adhesion of e-beam evaporated Ni/Al films in comparison to sputtered Ni/Al films as sputtered films may be deposited a few monolayers into the substrate. This is usually advantageous in terms of film adhesion. However, an energetic flux of sputtered material that is buried into the thin device oxide layer may mean that the oxide film is damaged, outweighing any adhesion advantage. It was decided that e-beam evaporation would be used to deposit the top electrode, regardless of the adhesion difficulties associated with it, while sputtering was still used to pattern the bottom electrode.

There was no lift-off step in this run as the thermal stress in the e-beam evaporated metal caused the metal to curl and peel away at the sites covered by photoresist but still adhere at the desired sites, leaving well defined devices (with a few exceptions). Knowing of the poor adhesion of the e-beam evaporated film it was chosen to leave the photoresist on the wafer, rather than dissolving it in acetone. The devices were then tested for its D.C. *i-v* characteristics with the photoresist still on the wafer. The results are summarized in Table 4. The cause of the poor adhesion was ultimately deemed to be oil backstreaming from the diffusion pump, as evidenced from the poor pumping capability and the oil residue on top of the gate valve. The solution, a retrofit of the pumping system to a cryogenic pump, came too late for this research.

Table 4 Summary of devices in highest yield wafer

Total Number of Device Sites	**78**	**%**
Malformed Devices	7	**9**
Shorted Devices	21	**26.9**
Open Devices	12	**15.4**
Highly Resistive Linear Devices	10	**12.8**
Functional Devices	28	**35.9**

The linear devices had a high resistivity, similar to the zero-bias resistance of the non-linear devices, and were most likely non-linear at much higher bias (or would be if they did not suffer dielectric breakdown). The malformed devices were devices that had not adhered or were in some way physically defective such that a tunnel contact was not possible. Open devices showed no current at any voltage and the shorted devices had a high conductance (on the same order as the designed shorted test devices in row 1). Full *i-v* characteristics of seven of the functional devices were immediately recorded (and not more since it was impossible to digitally transfer data to a computer since the SPA was equipped with an HP proprietary, and now obsolete, floppy drive format, so all data was transcribed by hand from listed values (~ 56-160 data points each measurement). Future measurements may be facilitated by interfacing with the HP-IB port on the SPA. The remaining devices were simply noted for their characteristics (shorted, open, malformed, etc.).

6.3.4 Run 4

Another wafer of devices was desired to confirm that the process used on run 3 was repeatable and to test whether the devices can withstand photoresist dissolution and subsequent cleaning with DI water. Run 4 had no devices physically adhering, despite using similar deposition parameters to the previous successful run. The significant difference was that the wafer was allowed to cool overnight in the chamber before vacuum was broken and the wafer was removed. Unlike the previous run, no curling was evident and lift-off by dissolution of the photoresist proved difficult, subsequently washing away the top layer of deposited metal along with the photoresist and confirming suspicions that the prior run would have not had properly formed devices if the photoresist was dissolved in acetone. Because the film was allowed to cool under vacuum it had less thermal stress and was therefore less likely to curl.

6.3.5 Run 5

To alleviate the problem of adhesion it was decided to use a higher deposition rate of Ni such that the evaporant flux is more energetic, and hence the resultant films have higher adhesion to the substrate than films deposited with less energetic flux. The process parameters were kept the same as in run 3 with the exception of deposition rate of the second electrode. However, Run 5 resulted in no functional devices (all shorts). Lift-off was performed by dissolution of the HPR 504 photoresist in acetone and of the LOR 5B resist by dissolution in photoresist stripper (Shipley 1165 available from Microchem). Surprisingly, metal curling was not observed on removal of the sample in run 5 *immediately* after deposition, however lift-off was more successful than run 4. The devices were tested at D.C. prior to removal of the LOR 5B and all sites showed 'short' characteristics (mA current at 0.1V bias). While it is not generally possible to determine the presence of pin-holes explicitly it can still be concluded that they are the most likely source of device failure for run 5 (the pin-holes are beneath a metal electrode. TEM can be used to 'search' for pin-holes but it is not feasible to check the entire junction area in this manner). From the deposition rates given in Table 3 a possible cause of the pin-holes can be inferred from the increased deposition rate of Ni between the successful run 3 and unsuccessful run 5. The 50% larger deposition rate was used in an attempt to increase adhesion (which succeeded in this respect).

While the e-beam deposited second electrode devices yielded in better quality diodes they also suffered from poorer adhesion than the sputtered devices. The more 'gentle' deposition by e-beam evaporation, as compared to sputtering, seems to result in a trade-off between adhesion and junction quality. Better adhesion between layers is achieved by using a more energetic flux. The better adhesion is most likely the result of incident particles being buried in the substrate and MIM oxide layers, thereby gripping mechanically into the oxides, while at the same time sacrificing the MIM's oxide uniformity and possibly damaging it completely. A better approach may be to find a suitable metal combination that will bind to the substrate and to the oxides and each other by intrinsic chemical attraction, rather than the mechanical adhesion employed thus far. Metals such as chromium and titanium could be used as adhesion layers, rather than the aluminum layer currently used. These metals are known for their adhesion properties and are commonly used as adhesion layers. Further, these metals readily grow native oxides, so a metallic short across the junction due to the adhesion layer is not likely.

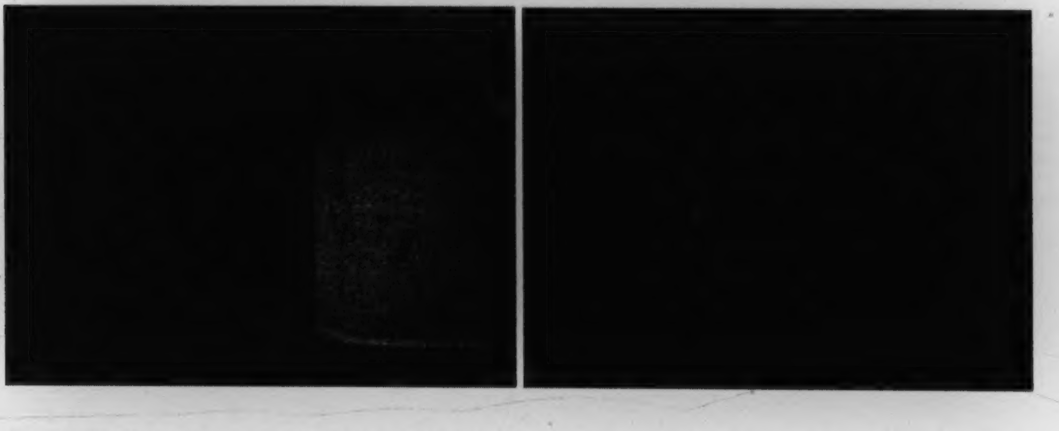


Figure 37 Diodes from run 5 (with the LOR resist layer unremoved). A malformed device (left) has had its top Ni layer removed. A diode with well-defined features (right)

6.4 Fits of Theory with Measured Data

Using MatLab® and a simplex based minimization function a theoretical fit of the measured data to Simmons' symmetric junction equations was made (MatLab code can be found in Appendix E). Simmons' medium voltage model (Equation 14) was used for $V < \phi$ and the high voltage model (Equation 15) was used for $V > \phi$, where ϕ was treated as a free fitting parameter. The fitting to Simmons' theory was difficult for some devices. The initial parameters had a large effect on how well the final fit correlated with the data, suggesting that the minimization function used did not consistently find the global minimum given a set of various initial guesses at the parameters. Due to the complexity of the function and the number of free parameters this non-ideal performance is deemed adequate. Some patterns do emerge and conclusions can be drawn from the fits. First, the work function difference between the nickel and nickel oxide is approximately 0.1-0.2eV. The theoretical work-function difference between nickel and nickel oxide is 1.3eV(ϕ_{Ni}≈5eV[87], ϕ_{NiO}=3.7eV [104]) so these value are a bit unexpected but given that thin-films are used (for which the work-functions may vary from bulk values) and that the thin barriers effectively lower the barrier height (due to image forces) these values are not unreasonable. Second, the tunnel barrier is smaller than corresponding average ellipsometrically measured values, suggesting an uneven insulator thickness and that conduction is dominated by current through hot-spots. These results and conclusions are similar to those found by several researchers.[22,105] Sample theoretical curve fits to data are shown in Figure 38.

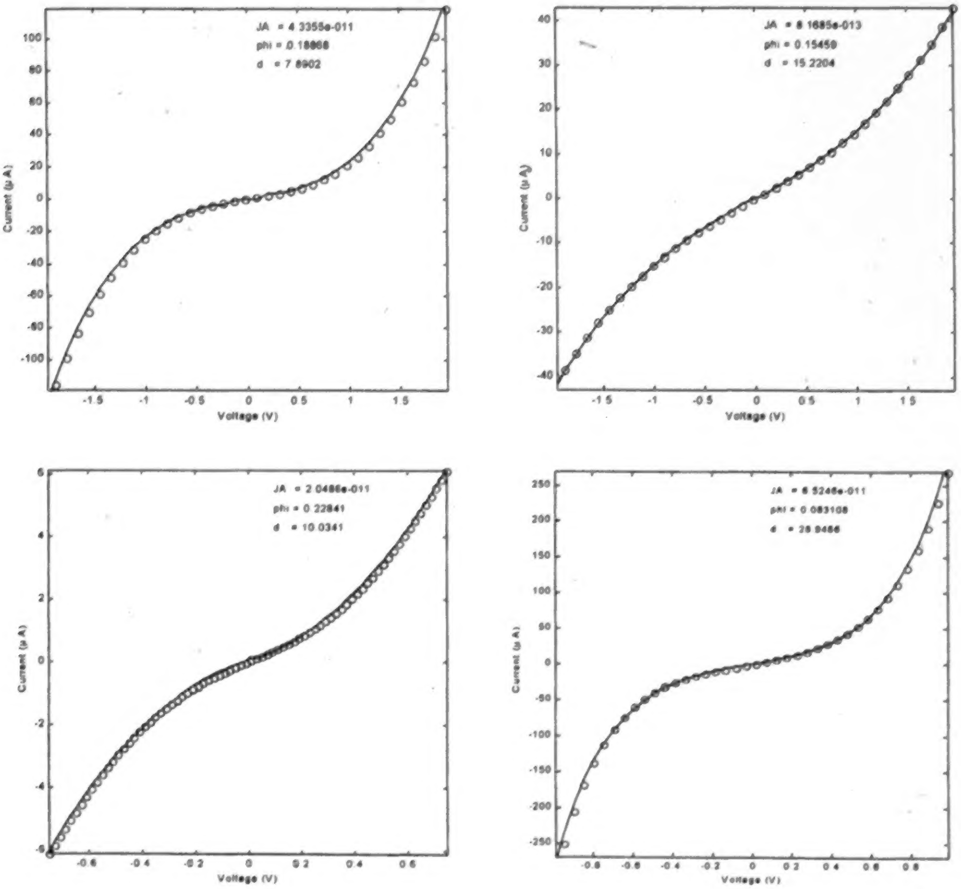


Figure 38 Selected fits of data with the theory of Simmons JA= junction area in m^2, phi = work function difference between metal and barrier in eV, d = barrier length in Å (clockwise from top left diodes are 7M, 7L, 7E and 6M)

For comparison, a curve fit by Scherer to a Cr/CrOx/Cr tunnel junction is shown in Figure 39. Considering that the measurement was taken at 4.2K and their junction fits are ill fitting at lower voltages the fits presented in Figure 38 are, at the very least, reasonable. Scherer's results also found a discrepancy between geometric and effective junction areas, but their ratio effective junction areas to geometric junction areas was only 10%-20%.

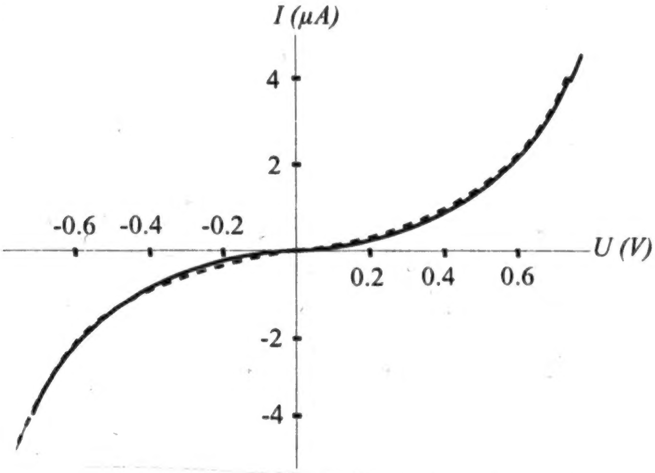


Figure 39 A curve fit using Simmons' theory by Scherer *et al*[31] of a Cr/CrOx/Cr tunnel junction

Keeping any of the four free parameters constant has led to even worse fits to tunnel junction models, as evidenced by works of Luo *et al*[106] and the work of the ITN group. However, very good fits are generally found for slightly non-linear devices. Highly non-linear device fits often have high residuals at the low or the high bias ranges. A similar condition is found in this research.

6.5 Summary

The major issue in the development of high quality tunnel junctions is the quality of the thin dielectric layer and presence of pinholes in lower quality oxides. The pinholes generally make the devices less non-linear (as in run 1) or cause the devices to fail entirely and can be minimized by use of a less energetic deposition technique (i.e. e-beam evaporation). Deposition by e-beam evaporation currently comes at a cost of poorer adhesion than sputtered films. In the future, high-quality pinhole free junctions that are constructed with films that adhere well may be possible with the e-beam evaporation of other thin adhesion layers, such as chromium or titanium. E-beam evaporation deposition rates should be kept low to minimize the energy of the evaporant flux and the base pressure should be kept below 10^{-6}torr to minimize reaction of the evaporant with remnant oxygen (a lower base pressure is preferred to minimize this effect especially due to the low rate of deposition).

7

Chapter 7: Tunnelling Transistor Design

7.1 Chapter Overview

A process for the fabrication of a metal-insulator tunnel transistor is proposed. Two process alternatives are presented for a critical process step. Design considerations for the transistor are presented and estimated parameters are given.

7.2 Transistor Design

The MITT transistors fabricated to date by other research groups have one defining drawback that is addressed here. Their processes for growing the tunnel junction insulator (TiO) have consisted of either local oxidation by anodization with an STM tip[39] or by anodization in an ethylene glycol dissolved ammonium tetraborate solution and an applied voltage (pictured in Figure 10 and Figure 40, respectively)[40]. In either case, the oxidation method does not lend itself well to fabricating large numbers of the devices in parallel; hence, the processes are not suitable for non-academic purposes. Nevertheless, the concept of a tunnelling transistor is very promising since it can be made with surface micromachining techniques, such that 3D circuits are possible since there is no requirement for a single crystal substrate. A proof of concept, fabricated by others, has already shown that the device functions (albeit not ideally) at room temperature.[39]

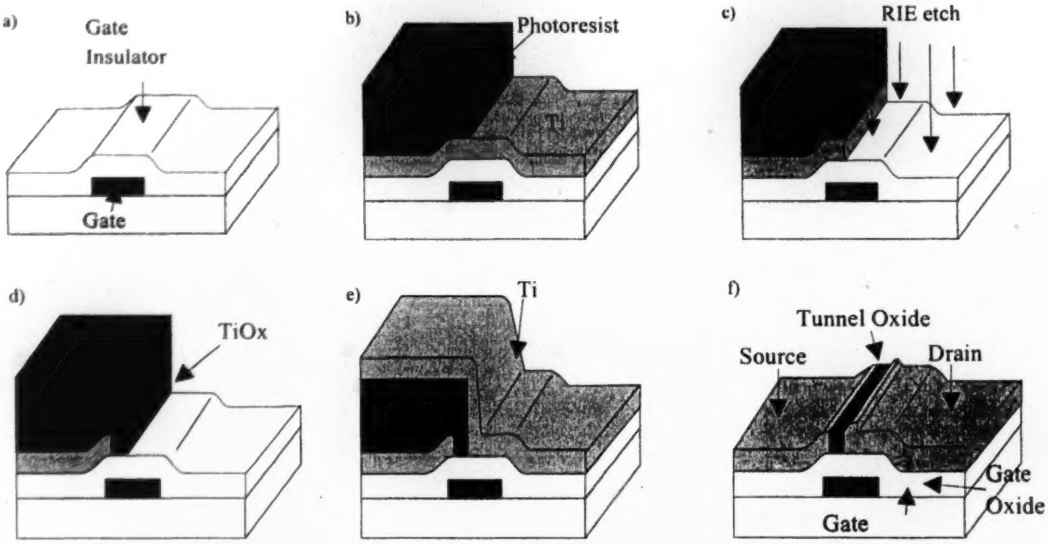
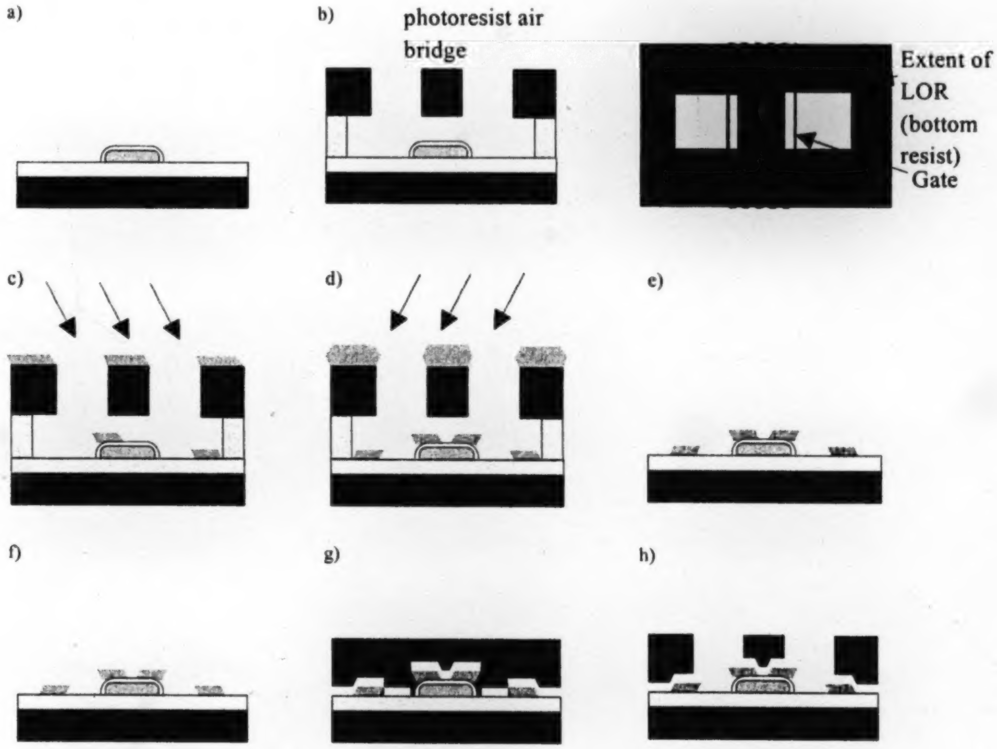


Figure 40 Process steps for an MITT fabricated with anodic oxidation in ethyline glycol-dissolved ammonium tetraborate a) Formation of gate insulator b) Ti evaporation and lithography c) Dry etching of half part of Ti d) Anodic oxidation of Ti e) Evaporation of Ti f) Lift-off process[40]

The fabrication process of the MITT, presented here, shares many similarities with the process of the MIM diode. The differences stem from the addition of a third terminal, a subsequent second tunnel junction, angled deposition and slightly altered geometries. However, the critical process step (i.e. oxide growth control of the gate and of the tunnel junction in one of the alternative processes) is essentially the same.

Using the deposition methods and the oxidation techniques of this research, the following process (process MITT A) may be all that is necessary to produce entire circuits in parallel. First, the gate metal is deposited by sputtering or e-beam evaporation. The gate oxide is then grown by the sputter-etch technique to a thickness of a few nm (2-5) (Figure 41 a). The wafer is then patterned with a bi-layer resist for the source and drain electrodes (a line that is perpendicular to the gate electrode, with a discontinuity, or break, in the vicinity of the gate. The lift-off resist layer (bottom resist) is completely removed in the vicinity of the gate such that the top resist layer forms a bridge along the gate (Figure 41 b). The source and drain are then successively evaporated (Figure 41 c and d), each at an angle that will deposit the source-drain electrodes close together above the gate. The photoresist is then stripped and the tunnel barrier is deposited (Figure 41 e and f, respectively). After etching away the barrier material from unwanted areas and filling in the interconnects the device is finished (Figure 41 g-k).



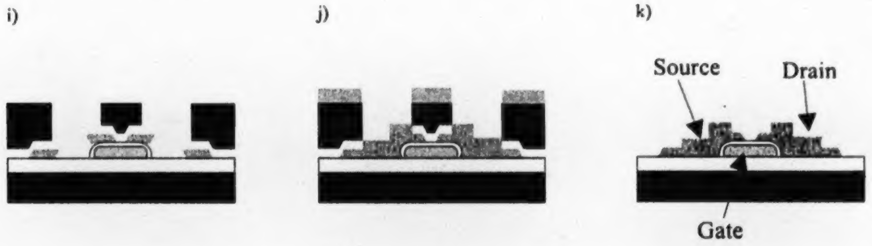

i)
j)
k)
Source
Drain
Gate

Figure 41 MITT process flow for a deposited tunnel barrier.

Alternatively, if it is determined that a high quality tunnel junction can not be fabricated with deposition techniques (and annealing, if necessary) then plasma oxidation can also be used in a slightly modified process (process MITT B). The source-drain electrodes can be deposited at angles such that they slightly overlap (Figure 42).

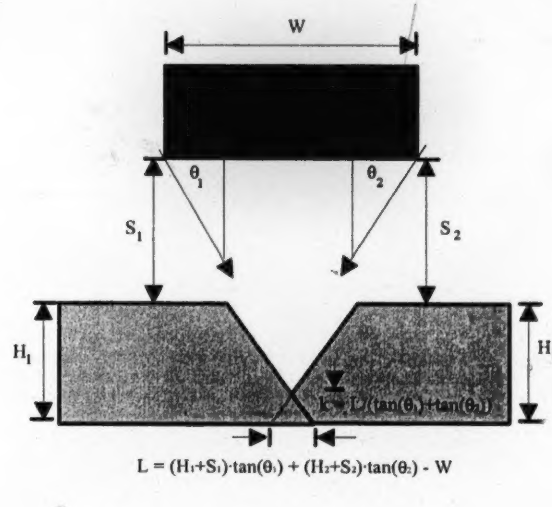

W
θ_1
θ_2
S_1
S_2
H_1
H_2
$k = L/(\tan(\theta_1)+\tan(\theta_2))$
$L = (H_1+S_1)\cdot\tan(\theta_1) + (H_2+S_2)\cdot\tan(\theta_2) - W$

Figure 42 Geometry of the overlapping source-drain tunnel junction method

After stripping the photoresist, a plasma oxidation technique can be used to oxidize the surface of the device such that at its thinnest point (the overlapping source drain region) it becomes completely oxidized and a tunnel barrier is fabricated (Figure 43). While this is a simplified case using only geometric effects and neglecting tapering and film growth effects, it serves to present the point. For it to be the thinnest point, the overlapping region L can only consist of the sloped portion of each electrode and for good separation of tunnel junction from metal electrode the film thicknesses H_1 and H_2 should be comparable (i.e. 20-50nm each). The overlapped region should not exceed ½ the length of the sloped region. In this way it can be ensured that the overlapping region is much thinner than the rest of the electrode (a necessary condition for the tunnel junction region to be completely oxidized, while the rest of the electrode has sufficient unoxidized metal for conduction). Prior works on the MITT have used or proposed 16nm to 30nm thick tunnel junctions.[40,41] Using these lengths as a guideline, with the 30nm junctions corresponding to lower leakage current at room temperature than the 16nm junctions, a target tunnel junction length of 20-40nm is chosen. Assuming that the overlap region length corresponds to the eventual tunnel junction length, the tunnel junction length can be prescribed. For instance, with film thicknesses of 20nm, a 45° angle and a 1μm air bridge that is 510nm above the tunnel junction, the overlapping region should be approximately 21nm +/- 1nm long and 10.5nm high. With a 21nm thick oxide the overlapping region should be entirely oxidized and a tunnel junction length should be 21nm near the gate oxide.

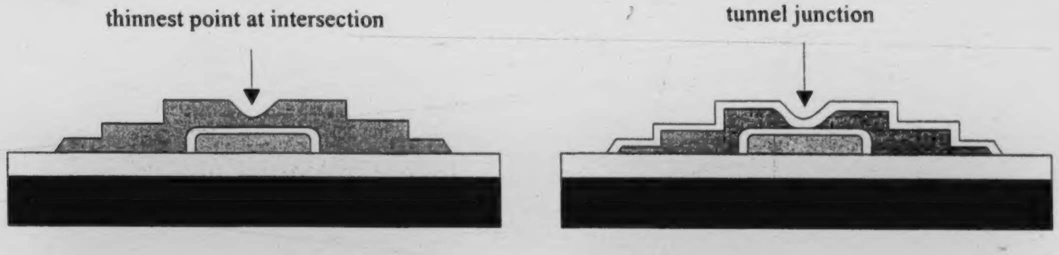


Figure 43 Slightly overlapping electrodes by angled evaporation (left) and plasma oxidation creation of tunnel junction (right)

The double-angle evaporation technique has the advantage that the source drain electrodes have sloped edges in the tunnel barrier region. By geometry, the shortest tunnel barrier is directly adjacent to the gate oxide, where the tunnel barrier is most affected by the applied gate voltage (the electric field distribution is more distorted near the gate oxide, as shown later in this chapter). The tunnel barrier is then perfectly positioned for maximum effect.

However, given that the evaporation is not uniform across the substrate (in terms of incident angle), a certain amount of variability across a wafer is expected. Given that the desired overlap between source and drain is just a few tens of nanometres and the geometric variables (air bridge width, height) are a few microns in dimension a fair amount of precision is necessary. The lithography must be carefully controlled for predictable results. Fortunately, this is something that can be controlled to a degree. A non-uniformity effect that is inherent in the process is the incident angle of the evaporant flux varying from one side of the wafer to the other (Figure 44). This effect can be partially alleviated by rotating the substrate 180° between the source and drain depositions such that the overlap is the same at points equidistant from the centre of the wafer and it is not as important effect as the lithography variation. The variables that can be changed are the substrate angle θ and the air bridge height (photoresist thickness) and width (developed feature size). Table 5 gives a summary of the variable effects for a typical system with L=30cm, W=10cm.

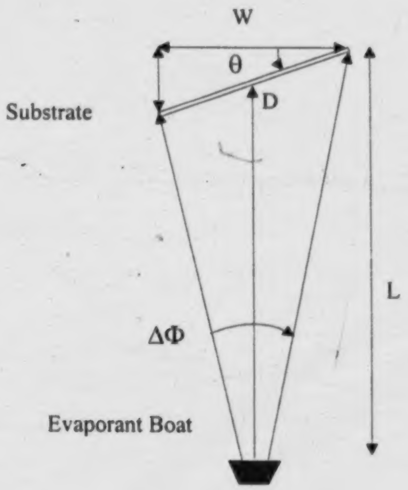


Figure 44 Evaporation angle non-uniformity (ΔΦ)

Table 5 Angled evaporation effect of variables on tunnel barrier geometry

Substrate Angle θ	Air Bridge Height H+S	Air Bridge Width W	Minimum Tunnel Barrier Length		Sensitivity of Angle Variation at Edge of Wafer (ΔBarrierLength in nm/ ΔAngle about Centre Angle)		Sensitivity to Lithography Variation at Wafer Edge (ΔBarrierLength in nm/ΔW)	
(deg.)	(µm)	(µm)	Centre (nm)	Edge (nm)	(nm /-1°)	(nm /+1°)	(nm /100nm)	(nm /100nm)
30	0.90	1	39.23	42.15	-41.46	42.31	-94.74	+115.8
30	1.75	2	20.73	26.41	-80.62	82.26	-96.50	+106.65
35	0.73	1	22. 30	24.60	-37.5	38.43	-93.15	+113.84
35	1.44	2	16.60	21.13	-73.98	75.81	-96.24	+106.38
40	0.61	1	23.70	25.54	-35.75	36.81	-93.23	+113.95
40	1.20	2	13.84	17.46	-70.32	72.41	-96.07	+106.18
45	0.51	1	20.00	21.45	-34.98	36.22	-92.86	+113.49
45	1.01	2	20.00	22.87	-69.27	71.73	-96.33	+106.47
50	0.43	1	24.91	26.04	-35.57	37.08	-93.28	+114.00
50	0.85	2	25.98	28.22	-70.32	73.31	-96.58	+106.75

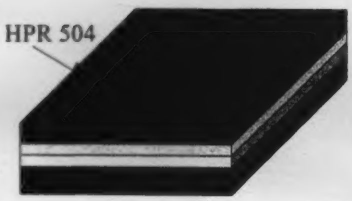
As is clear from the above table, the challenge is not so much to control the variation of the tunnel barrier length as a function of position on the wafer, but rather to control the lithography parameters (feature size and photoresist thickness). Since it is expected that the lithography will be harder to control than the angle variation of the substrate holder it is desired to minimize the sensitivity of the lithography variation and therefore choose a high incident angle θ (close to 45°). A very high degree of precision is necessary for this process to succeed, particularly in the lithography.

To more accurately predict the thin film profile the deposition can be simulated using a thin-film deposition software simulator, SIMBAD, developed at the University of Alberta, and used in industry.

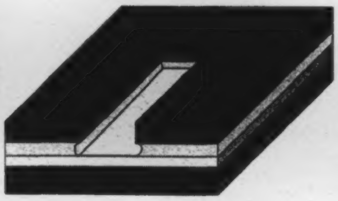
7.3 Process Steps

A step-by-step process follows. The process flow is specifically for the deposited oxide source-drain junction process (MITT A), though the plasma oxidation process (MITT B) differs only by the geometry of the angled evaporation, discussed previously and by the oxidation of the source-drain tunnel barrier in place of its deposition.

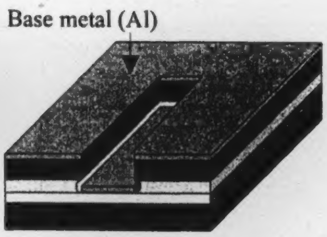
1



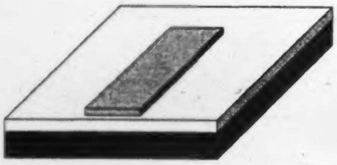
2



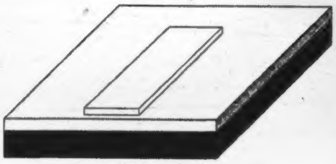
3



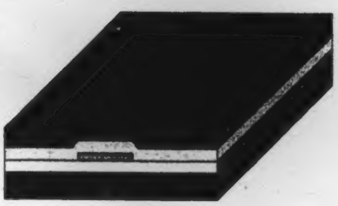
4

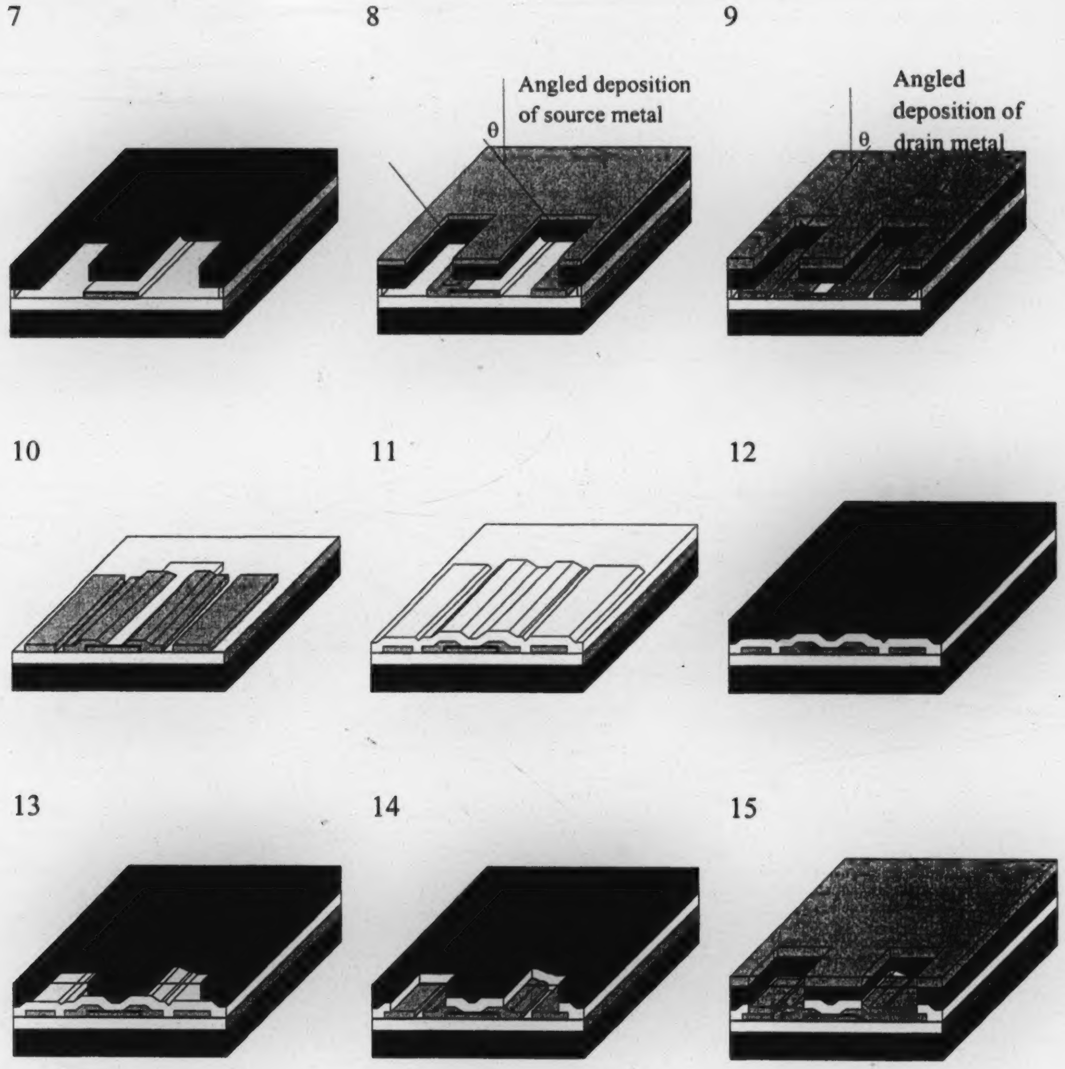


5

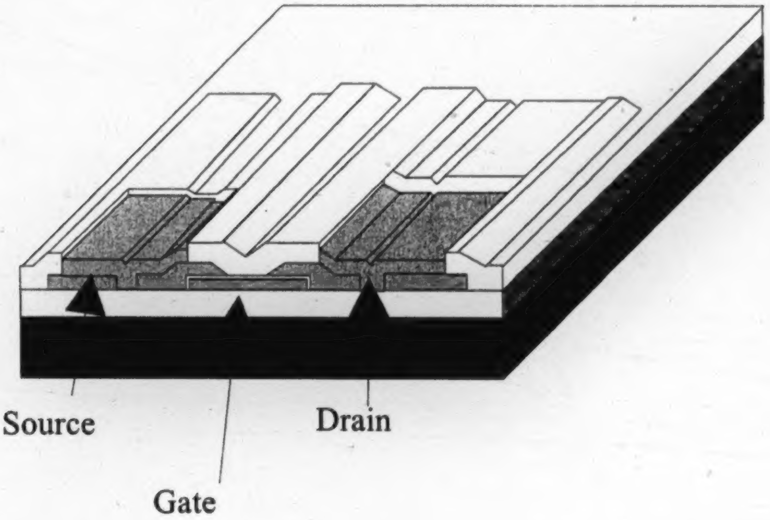


6




Angled deposition of source metal
θ
Angled deposition of drain metal
θ

16



As is quickly apparent, the proposed MITT process shares similarities with the MIM diode process. The major difference comes in the angled evaporation of the source and drain electrodes and reactive sputtering of the tunnel junction. The details of the process, that are the same as in the MIM diode processes, are omitted for clarity.

1. The process begins with a photoresist covered Si/SiO_2 wafer, prepared in the same manner as in the MIM diode processes.

2. A thin line (2μm) line is etched in the photoresist by exposure to a gate pattern mask.

3. Gate metal is deposited through either e-beam evaporation or sputtering. Aluminum is a good preliminary choice of metals as it adheres well to SiO_2, and more importantly, has a large work-function difference with its oxide.

4. The photoresist layers are stripped in acetone and photoresist stripper.

5. The gate is oxidized using the plasma oxidation technique presented in this research.

6. The bi-layer of photoresists is spun. A thicker lift-off resist layer may be necessary than previously used (400nm) since the double-angle evaporation must be such that the electrodes deposited are close together and a high incident angle evaporation may be undesirable. If 2μm lithography is the minimum achievable reliably (for pattern etch and alignment purposes), then an LOR 7B or 10B or another thicker resist may be desirable.

7. The source-drain electrode patterns are etched in the photoresist stack.

8. The source is e-beam evaporated to a thickness of a few 10s nm (20-50), at an incident angle $\theta_{dep\text{-}drain}$, normal to the substrate and in a direction towards the gate electrode. Nickel is the first choice of source-drain metal, for reasons presented later.

9. The drain is evaporated at an angle $\theta_{dep\text{-}junction}$, directed towards the gate electrode. The angle is measured from the normal to the substrate and is the same angle as in step 9, but rotated 180° about the plane of the substrate.

10. Photoresist is stripped off the wafer.

11. The tunnel junction is reactively evaporated (evaporated in a rich oxygen atmosphere so that all incident evaporant is an oxide). If necessary, an annealing step follows to make the oxide more uniform in terms of crystallinity.

12. The photoresists are spun as before.

13. The photoresist is patterned for areas where it is desired to remove the deposited oxide.

14. A chemical or plasma etch is applied (a fluorine based chemistry for most metal oxides, such as BOE chemical etch or CF_4 plasma etch). For metal oxides that are not efficiently etched by fluorine, as is the case with NiO, an Ar^+ sputter etch can be applied.

15. A thicker layer of source-drain metal is deposited for pads and interconnects.

16. The device is complete.

For alignment purposes, the dimensions of the device are kept non-minimal. The gate should be at least 2µm in width (length in CMOS terminology) such that a 1µm mask aligner can, with some certainty, align the source-drain ends onto the gate.

7.4 MITT Materials and Geometry Considerations

One of the challenges involved in the construction of an adequate device (capable of room temperature operation) involves the proper choice of materials. For the conduction to be mainly source-drain (and not gate-drain or gate-source) the tunnel barrier height at the source drain junction should be low, but not so low as to allow high conductivity in the off state (i.e. low enough to reach the F-N regime with a reasonable bias but high enough so that the F-N regime is not reached with a low bias). The gate insulator, by contrast, should have a high barrier height (2V or greater[38]) so as to limit current through the barrier. The proper choice of materials that satisfy these requirements involves proper interpretation of data from the MIM diode (extraction of device parameters from graphs of data). A preliminary choice for the gate and gate oxide is an

Al electrode covered with a grown oxide of Al_2O_3 ($\phi_B = 2.0V$). Nickel is a good choice due to the low barrier height of the Ni-NiO contact (0.1-0.2eV found in the curve fits of this research), despite the work-function difference between nickel (5.15eV) and aluminum (4.28eV). The barrier between Al_2O_3 and nickel should then be approximately 0.87eV, which is still greater than the 0.2eV Ni-NiO work-function difference.

Another challenge is to optimize the geometry of the device, for similar reasons (i.e. high on state current, low off state current). The gate insulator must be thin (a few nm) so that it can be grown using the plasma oxidation method used for the MIM diodes of this thesis. The gate junction is critical such that little tunnel current flows from gate to source or drain while the electric field is sufficient to distort the tunnel barrier. The physical height of the tunnel junction oxide should be kept small such that the tunnel barrier has its potential profile significantly affected by the gate voltage (i.e. if the tunnel junction height was large the effect of the gate voltage on the tunnel insulator would be minimal as much of the tunnelling would occur far from the gate).

The electrostatic characteristics of an ideal MITT device with gate thickness of 5nm and relative permittivity of 11.5 (Al_2O_3), source-drain thicknesses of 10nm and tunnel insulator length of 20nm and relative permittivity of 4 (NiO) are shown in Figure 45. Simulations with the finite-element analysis program ANSYS®/ED Multiphysics v6.1[107] (on Windows 98 based PC) were run for gate voltages of 0V, 2.5V and 5V and 5V between the source and drain (see Appendix E for ANSYS code). Through electrostatic simulation the shape of the tunnel barrier can be determined since the potential distribution in the tunnel oxide corresponds to the potential distribution (and so electrical shape) of the conduction band of the insulator in a band diagram. The tunnel junction barrier is deformed (electrically) due to the varying electric field distribution, most notably in the region directly adjacent to the gate oxide. The source, drain and gate are omitted from the simulation since the electric field in the terminals is assumed to be constant (i.e. 0). When a low gate potential is applied (Figure 45 b)) a large change in potential occurs near the drain electrode at 5V, while a large part of the tunnel region is at 0V (high electron potential). This implies that the shape of the Fowler-Nordheim region is convex with an electrically thick barrier. Conversely, when a potential of 5V is applied to the gate the large potential change occurs near the source at 0V and a large portion of the tunnel barrier is at 5V (low electron potential). This implies that the Fowler-Nordheim regime is concave and electrically thin.

The choice of aluminum oxide for the gate barrier is further justified by its high dielectric permittivity, which increases its effect on the tunnel barrier in comparison to a gate oxide with a lower dielectric permittivity (Figure 45 e)). By similar reasoning (i.e. better penetration of gate electric field) thinning of the gate oxide also should result in better gate control over the tunnel barrier shape (and therefore source-drain tunnel current).

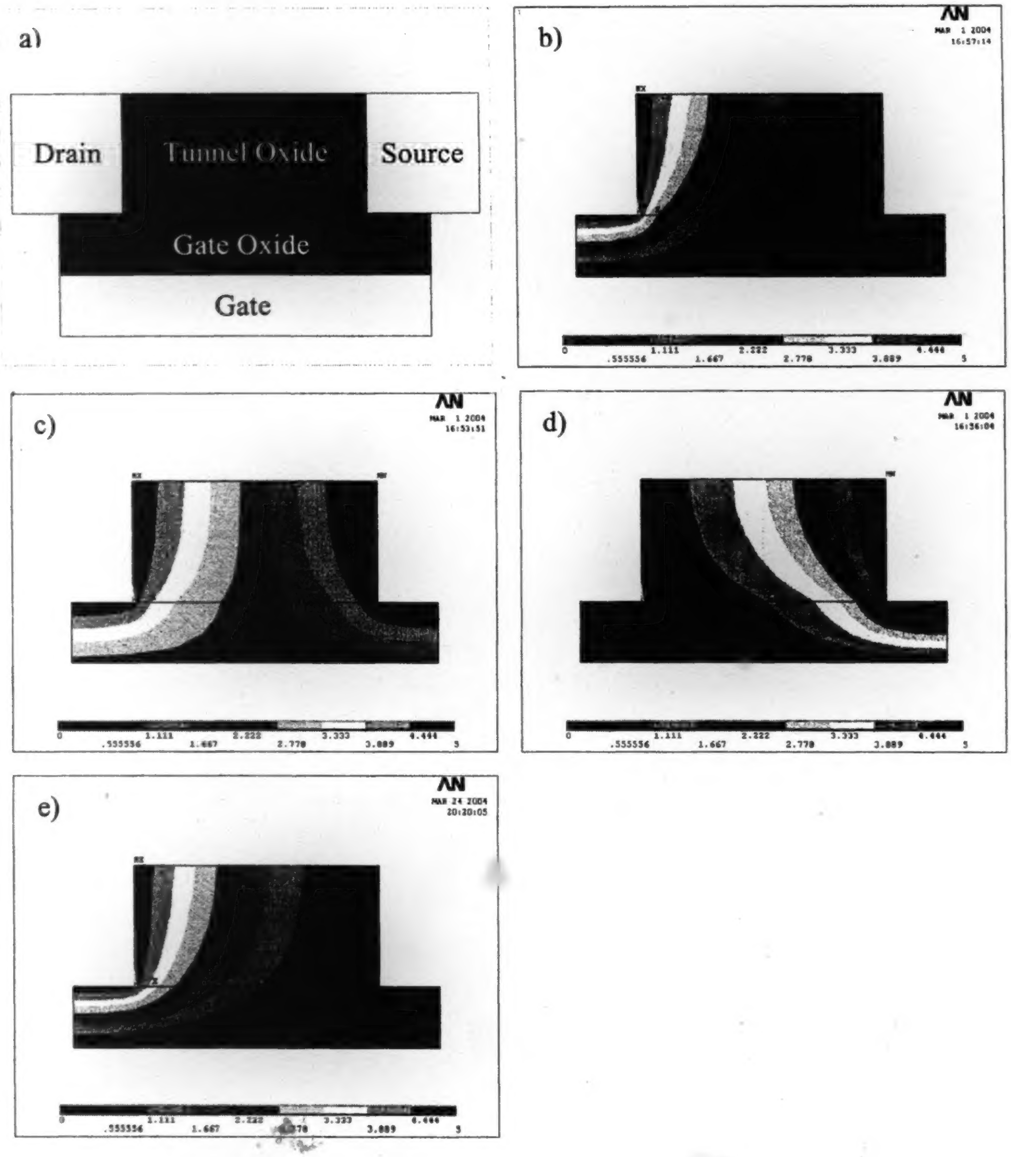

a)
Drain
Tunnel Oxide
Source
Gate Oxide
Gate
b)
AN
MAR 1 2004
16:57:14
c)
AN
MAR 1 2004
16:53:51
d)
AN
MAR 1 2004
16:56:04
e)
AN
MAR 24 2004
20:30:05
0
.555556
1.111
1.667
2.222
2.778
3.333
3.889
4.444
5

Figure 45 Finite-element simulation of the electrostatic characteristics of an ideal straight-walled MITT, showing deformed field in the tunnel region with applied gate bias. a) cross-section of MITT, b) gate voltage = 0V, c) gate voltage = 2.5V , d) gate voltage = 5V, e) gate voltage = 0V and relative dielectric permittivity of gate lowered to 4.

While further research is necessary into optimal material choices such that this tunnelling is minimized (high ϕ_B) the fabrication of the gate is consistent with the fabrication method of the tunnel junctions of this research. Therefore the fabrication of an MITT should be possible with current techniques, assuming the adhesion issues of the present work are adequately addressed. Given that the metals prescribed for MITT fabrication have high intrinsic adhesion with substrates the adhesion should not present much of a problem if clean sputtering or e-beam evaporation systems are used at low base pressure.

8

Chapter 8: R.F. Characterization and Harmonic Radar Tag Design

8.1 Chapter Overview

This chapter presents a harmonic radar tag design incorporating the MIM diode of the present work. The ultimate goal of this project is to integrate a smart-dust system utilizing a harmonic radar communication scheme. Ultimately, this could be comprised of on-chip MEMS sensors, circuitry, rectification and the antenna-diode tag, at a high frequency (>10GHz) such that the entire system is on the order of a few square centimetres or less in area. Due to its high-impedance properties a folded-dipole antenna design is presented for immediate integration with the MIM diode at a lower frequency (up to 1GHz) due to measurement device limitations. R.F. spectrum scans of the MIM diode are first presented as evidence of MIM diode harmonic generation at R.F. frequencies.

8.2 R.F. Measurement of Diode Properties

The R.F. characteristics of the MIM diode have been measured by observing MIM diode generated harmonics at R.F. frequencies. The non-linear characteristics of the MIM diode at RF frequencies were measured with the spectrum analyzer set-up shown in

Figure 46. The equipment was made available by Brian Moore and Scanimetrics (Edmonton, AB, Canada) after R.F. equipment of the Electrical and Computer Engineering department was deemed unusable since, as a whole, it lacked several key components and attempts to access good R.F. equipment at local institutions proved fruitless. Originally, to avoid the problems of gaining access to adequate R.F. equipment, the testing plan was to use a non-contact method of testing the harmonic radar tags themselves with a Recco system interrogator. After it was discovered that the MIM diodes have a weaker first harmonic (on which the Recco system is based), relative to the second harmonic, it was decided to test the diodes directly and then design a folded dipole antenna to match to the diodes.

The spectrum analyzer used for testing the diodes was a Tektronix 3088 (0-3GHz; in practice, due to low frequency noise its usability begins at ~5MHz) and an HP 8702B lightwave component analyzer (in E/E mode) was used as the RF source (300kHz-3GHz, -15-20dBm; in practice only –3 to 9dBm was used). Due to the stray capacitances of the wafer probes the set-up was limited to measurements below 1GHz. The probes themselves (Alessi MD series micropositioners) were low frequency probes that were found to be usable only to 50MHz without the soldered grounding connections between the outer conductors of each probe. Above 50MHz the antenna properties of the poorly grounded probes bypassed the device under test and there was an increasingly strong (with frequency) over-the-air connection between the stimulus and measurement probes. The addition of the soldered ground connections moved this 'roll-off' frequency to 1GHz. SMA terminations were used throughout, except at the probes (BNC) and measurement equipment (N-type). The R.F. source (HP 8702B) fed the input signal to one terminal of the MIM diode through the wafer probes. The directional coupler was used to remove frequency dependent reflections by calibrating allowing only unidirectional signal flow and by calibrating the signal through its self-referred (-11dB) connection. The spectrum analyzer then measures the signal through the diode at the other diode terminal.

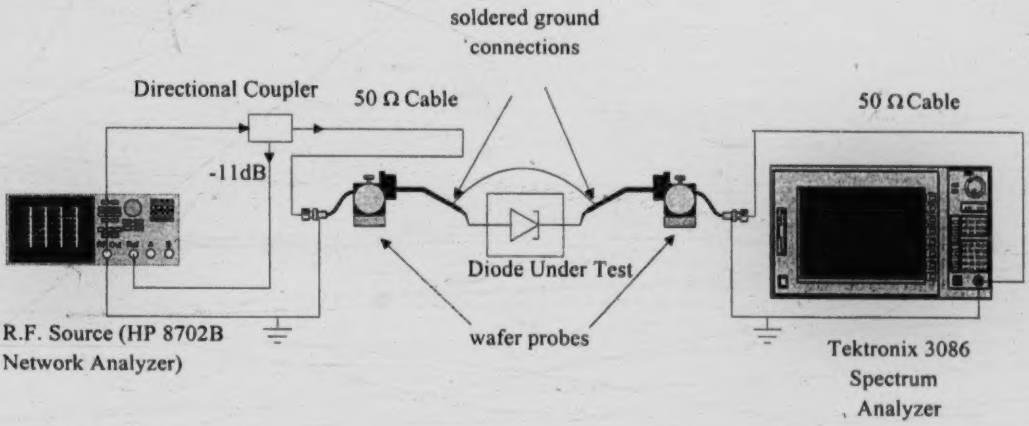

soldered ground connections
Directional Coupler
50 Ω Cable
50 Ω Cable
-11dB
Diode Under Test
R.F. Source (HP 8702B Network Analyzer)
wafer probes
Tektronix 3086 Spectrum Analyzer

Figure 46 Measurement of MIM diode RF characteristics.

The testing consists of simply applying a fundamental frequency (swept across a small range of successive frequencies) and observing the resulting signals at harmonic frequencies. The stimulus frequencies were swept across a small range of frequencies (about +/-10% of centre frequency) so that there is no confusion between the response of the diode and spurious or quiescent R.F. signals. This method of testing also directly shows the suitability of the diode in a harmonic radar system. The harmonic generation can be seen in Figure 47 and Figure 48. When an MIM diode is probed at sufficient input power (assuming a sufficiently non-linear *i-v* characteristic) the diode produces a current at the second harmonic frequency and a smaller current at the first harmonic, as expected from the D.C. *i-v* curves which, when expressed by polynomial expansion, show large second harmonic (3rd order) and smaller first harmonic (2nd order) terms. With no device and a short between the probes the pulses at the harmonic frequencies do not exist. The detection limit of the equipment (-70dBm) was low enough that this effect could clearly be seen and the spurious and quiescent noise is clearly distinct from the signal. Also, with an open circuit (no connection to a device) there are no significant artefacts that can be mistaken for a harmonic response. The conversion loss, defined as the difference in power level between the fundamental frequency and the harmonic frequency, varies by power level since the diode enters its non-linear region only after a turn-on voltage. The first and fourth harmonic first appear at an input power of –3dBm, while the second harmonic appears at a power of 0dBm. The harmonic response of the diode clearly increases with increasing input power and the conversion loss diminishes. The conversion losses at several frequencies and powers of –3dBm to 9dBm are summarized in Table 6.

Table 6 Conversion efficiency vs. input power for MIM devices at RF frequencies

Centre Frequency	Power Level	Conversion Loss (1st harmonic)	Conversion Loss (2nd harmonic)
10MHz	-3dBm	> 56dB *	49dB +/- 3dB
10MHz	0dBm	56dB +/- 2dB	43dB +/- 3dB
10MHz	3dBm	44dB +/- 1dB	42dB +/- 1dB
50MHz	0dBm	54dB +/- 2dB	42dB +/- 1dB
100MHz	6dBm	38dB +/- 3dB	22dB +/- 3dB
100MHz	9dBm	45dB +/- 1dB	24dB +/- 1dB
200MHz	0dBm	49dB +/- 2dB	39dB +/- 2dB

* Harmonic power level below detection limit of measuring device

These results show that the MIM device is a viable R.F. device for use in harmonic generation and mixing (as a consequence of the harmonic generating non-linearity). Despite the fragility of the devices (to ESD and over-voltage) and the challenges involved with the probe set-up, repeatable and reliable results (by virtue of the frequency sweep) were attained. The device with the best characteristics was found to function up to 600MHz (second harmonic of 200MHz) before it too failed by producing an open circuit across the diode electrodes. However, the device had been subjected to several prior measurements, including a sweep at 9dBm, prior to its failure. It is expected that these devices could function to beyond the 1GHz limit imposed by the measurement equipment, if not stressed by prior measurements at higher power (i.e. 9dBm).

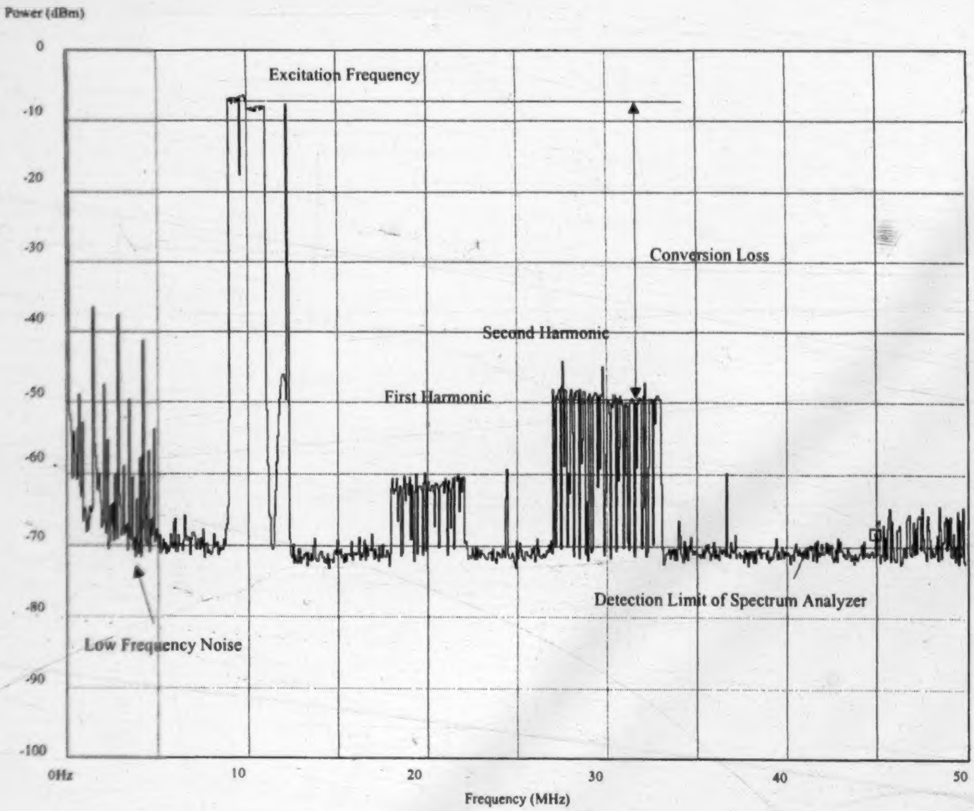


Figure 47 Spectrum scan of harmonic generation by the MIM diode at 9-11MHz excitation and 3dBm applied power.

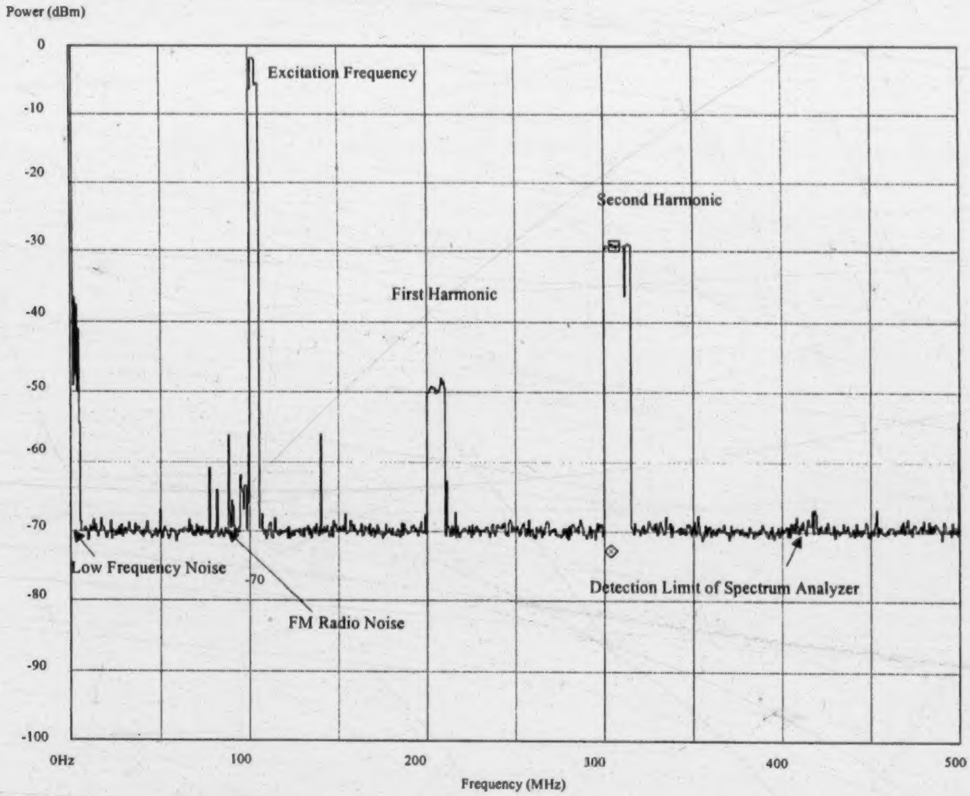


Figure 48 Harmonic generation at 100MHz excitation and 9dBm applied power

8.3 Tag Design

A harmonic radar tag is a relatively simple, yet practical and powerful, application for the tunnel diodes fabricated in this research. A simple system consists of just an antenna connected to a single MIM diode. A more complicated system could include enhanced R.F. properties by integration with R.F. MEMS or enhanced functionality by integration with MEMS based sensors. Because the MIM diode can potentially be responsive to very high frequencies the antennas can be made extremely small. This enhances the possibility of an integrated harmonic radar system, with or without MEMS sensors, that is smaller than a few millimetres on a side. The goal for the current design, though, is for construction of a lower frequency tag, such that construction and testing is more easily facilitated and to give a knowledge base from which future high frequency tags can be designed.

Prior work in harmonic radar tags, at the University of Alberta/Alberta Microelectronics Corporation, by Dr. Chris Backhouse and Brian Moore involved integrating an HP Schottky barrier diode (HP 2801 series) with a wire antenna. The tags were small enough that they were used effectively to track butterflies (by periodically checking for their presence at different locations) with a 30m range.[42] The rated dielectric breakdown of the diodes was ~80V but the diodes still had a problem with blowout. The problem was even worse for more conventional diodes (the HP 2801 diodes used were of military grade and have been discontinued) that had a dielectric breakdown of a few volts. A subsequent design involved integrating the diode with a lithographically defined folded dipole antenna on a PCB board (Figure 49). These were empirically found to work well with the Recco interrogator up to distances of 12m.[108]

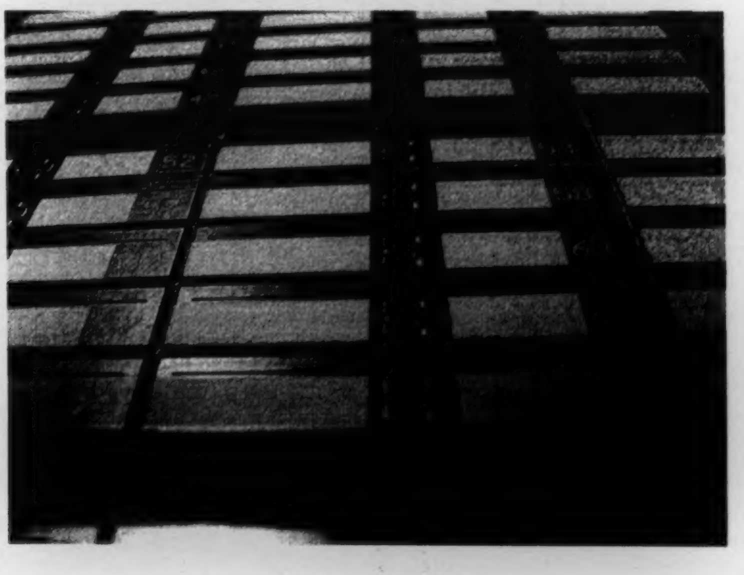


Figure 49 Unpopulated (PCB) R.F. tag by Moore/Backhouse

A significant problem with the wire tags was that they had to be individually and manually assembled, pushing the cost of the tags to about $50. Further, the HP diodes were themselves difficult to acquire (they are now obsolete). By constructing the tags using microfabrication techniques many devices could be built concurrently, without any need for manual assembly, greatly reducing the cost of each individual device and eliminating dependence on external supply. The current design involves just two systems: the antenna and the MIM diode.

8.4 Antenna Design

A conventional $\lambda/2$ dipole antenna has a terminal resistance of approximately 70Ω at its resonance frequency. When designing RF diode systems the criterion used to match the diode to the rest of the circuit is the diode's "video resistance." Simply put, the video resistance is the small signal, or differential resistance of the diode at the bias point. Since the MIM diodes presented here are unbiased their video resistance and zero bias resistance are the same at small ($V<\phi$) input voltage and similar at higher voltages. Considering the fact that the zero-bias resistance of most high quality (very non-linear) MIM diodes produced in this research and elsewhere is in the kΩ to hundreds of kΩ range the impedance mismatch between an MIM diode and a simple dipole could effectively prohibit its functionality (due to impedance mismatch power reflection). To improve coupling, a better impedance match is desired between the MIM diode and the antenna to which it is connected. This can be achieved through impedance matching techniques, but with each impedance matching element there may be an associated loss. It is desired to minimize the range of impedance between the antenna and the diode prior to final impedance matching. To this end, a high impedance antenna is desired.

An antenna that has a user-defined impedance is the loop antenna, one version of which is the folded λ/2 dipole. The folded dipole has a radiation resistance of 300Ω when the thickness of the outer conductor d_2 equals the thickness of the inner conductor d_1 (Figure 50). However, the antenna can be readily tuned over a wide range of impedances.

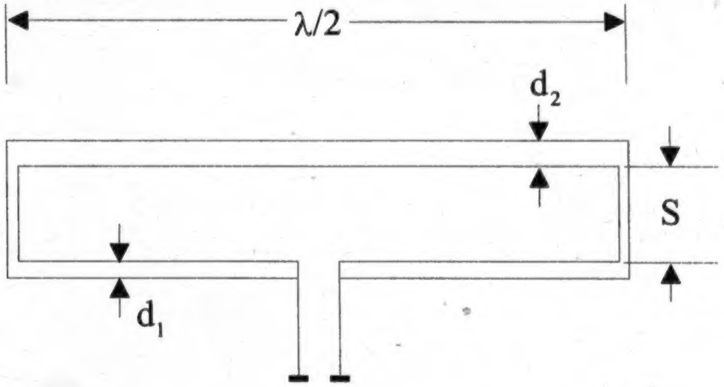


Figure 50 λ/2 folded dipole design parameters

The impedance at the terminals is close to purely resistive for this type of antenna[109] and its impedance can be changed by varying the thickness and spacing of the outer conductor relative to the inner conductor and by varying the spacing S relative to the conductor thickness d_2. For example, when d_2 is $3.5\cdot d_1$ and the conductor spacing S is $1.5\cdot d_2$ the impedance is approximately 3kΩ, a 10 × increase over the case of $d_1 = d_2$. The step up ratio over the $d_1 = d_2$ case is given by: [110]

$$r = \left(1 + \frac{\log\left(\frac{2S}{d_1}\right)}{\log\left(\frac{2S}{d_2}\right)}\right)^2 \tag{43}$$

for $S \gg d_1, d_2$.

The generalized impedance is given by: [110]

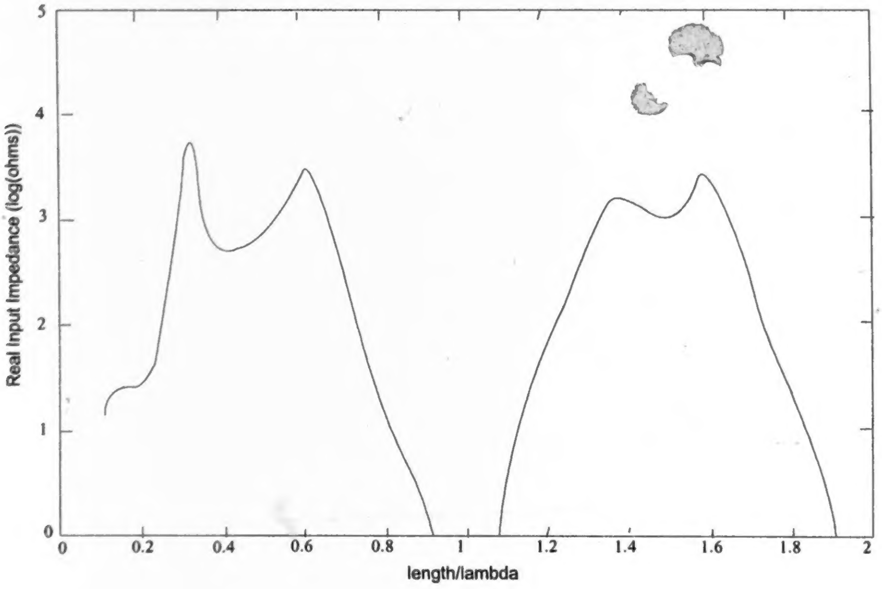
$$Z = \frac{2r \cdot Z_r Z_f}{r \cdot Z_r + 2Z_f} \tag{44}$$

where Z_r is the impedance at resonance, $Z_f = jZ_0\tan(kl)$, $k = 2\pi/\lambda$ and l is the length of the dipole and Z_0 is the characteristic impedance of the circuit, given by:[110]

$$Z_0 = 120 \cdot \cosh^{-1}\left(\frac{s}{2\sqrt{d_1 d_2}}\right) \tag{45}$$

for a free air-dipole. The microstrip characteristic impedance of the dipole is assumed to be the same as above, but divided by the square root of the effective dielectric constant (calculated below).

At off resonant frequencies, the folded dipole antenna drops significantly in impedance. Figure 51 shows the real and imaginary impedance for a folded dipole with the design dimensions given in Appendix D (the model itself is a complex function and can be found in the "Handbook of Antenna Design" by Rudge *et al.* and "Antenna Theory" by Balanis[110,111]). As can be seen in the graph of real impedance, R_{in}, the antenna is resonant at $\lambda/2$ and $3\lambda/2$, but has no impedance (and therefore no power coupling to the e-m wave) at antenna length of λ (fundamental). There exists a trade-off between antenna size, which could be made to be smaller than resonant, and impedance matching. Due to the large impedance of the diode it is not seemingly possible to match the antenna to the diode at off-resonant frequencies. However, a perfect match may not be necessary. The tags shown in Figure 49 functioned, despite the fact that the antenna dimensions were $\lambda/5$, rather than the resonant $\lambda/2$. In the near term, it is prudent to design the antenna for the resonant frequencies, as this represents the safest course of action, despite the large size of the antenna.



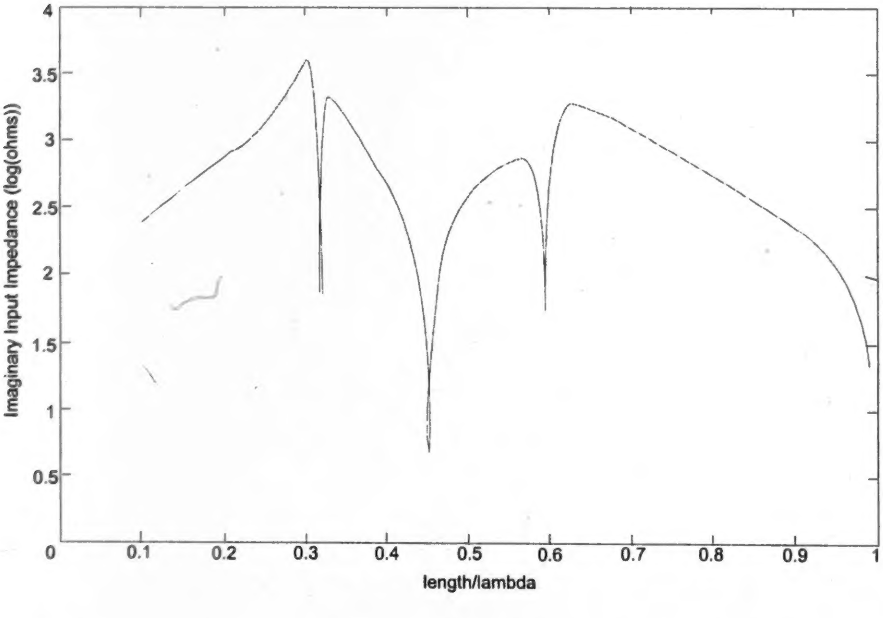


Figure 51 Input impedance (top-real, bottom-imaginary) of a folded dipole antenna

The MIM diode can be modeled as a simple resistance in parallel with a capacitance at frequencies up to the infrared.[112] Thus, the equivalent diode resistance, or video resistance, can be measured by measuring the insertion loss of the diode in a series resistive circuit at the frequency of interest.[113] A transmission line model is used as a basis for the design of the antenna. The antenna is treated as an asymmetric coplanar strip (CPS) transmission line on an insulating substrate (i.e. SiO_2), shown in Figure 52. The Widths W_2 and W_1 correspond to antenna dimensions d_2 and d_1. The insulating substrate thickness is taken to be large (>>W1, W2) since, as will be presented shortly, the substrate effect reduces the effective length of the antenna. A fundamental frequency of 2GHz is chosen so that the entire tag can be fit onto a 4” substrate. Given that this is a prototype design, the thick substrate can be made of bulk (i.e. non-microfabricated) material during testing. The silicon backside is assumed to be etched away, or a pure glass substrate is to be used, so that only SiO_2 remains as the substrate. Additional bulk material is assumed be added below the microfabricated wafer to increase the substrate height during testing to simulate the condition of an electrically thick substrate. Since the dimensions of the prototype design are fairly large (several centimetres) the effect of the substrate wafer alone is minimal and bulk material must be added to the bottom during testing. It is expected that, as the device is scaled to millimetre dimensions, the effect of the substrate wafer itself will be significant and therefore should be accounted for in the prototype design and testing, through simulation of the condition (i.e. keeping the scaling the substrate thickness along with the dimensions of the antenna).

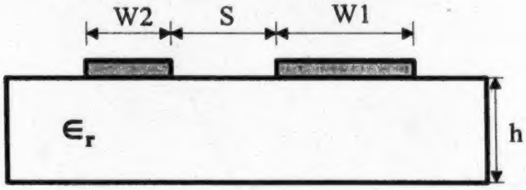


Figure 52 Asymmetric coplanar strip line (with a finite dielectric thickness) geometry

Considering the effect of the substrate, the effective wavelength of a received or transmitted signal at the antenna must be computed, and is given by:

$$\lambda_{eff} = \frac{\lambda}{\sqrt{\epsilon_{re}}} \tag{46}$$

where the effective dielectric constant, ϵ_{re}, is given by:[114]

$$\epsilon_{re} = 1 + q \cdot (\epsilon_r - 1) \tag{47}$$

where,

$$q = \frac{1}{2} \cdot \frac{K(k_4)}{K'(k_4)} \cdot \frac{K'(k_3)}{K(k_3)} \tag{48}$$

$$\frac{K(k)}{K'(k)} = \begin{cases} \dfrac{\pi}{\ln\left(\frac{2\left(1+\sqrt{k'}\right)}{\left(1-\sqrt{k'}\right)}\right)} & for \quad 0 < k \le \frac{1}{\sqrt{2}} \\[2ex] \dfrac{\ln\left(\frac{2\left(1+\sqrt{k}\right)}{\left(1-\sqrt{k}\right)}\right)}{\pi} & for \quad \frac{1}{\sqrt{2}} \le k \le 1 \end{cases} \tag{49}$$

$$k' = \sqrt{1-k^2} \tag{50}$$

$$k_3 = \sqrt{\frac{W_1}{W_1+S} \cdot \frac{W_2}{W_2+S}} \tag{51}$$

$$k_4 = \sqrt{\frac{\sinh\left(\pi W_1/2h\right)}{\sinh\left(\pi(W_1+S)/2h\right)} \cdot \frac{\sinh\left(\pi W_2/2h\right)}{\sinh\left(\pi(W_2+S)/2h\right)}} \tag{52}$$

where t is the thickness of the substrate and d is the conductor width and ϵ_r is the relative dielectric constant of the substrate. The layout of the final design is given in Appendix D, for an antenna with an input resistance of 1239Ω at first resonance and 2335Ω at second harmonic resonance at l=0.5036λ, with the resonances taking into account the diode reactance of a 1x3μm 40Å Ni/NiO/Ni diode. Table 7 shows the effect of the substrate on the antenna parameters at constant antenna dimensions (as a fraction of wavelength). Clearly, if the relative antenna dimensions are kept the same, the substrate has little effect on impedance, but a large effect on resonant length. The theoretical case of an alumina base which has a high ϵ_r, shows that it too needs a large base thickness to have an effect on the dipole length for resonance.

Table 7 Substrate effects on antenna parameters. Only electrically thick (relative to wavelength) substrates have effect on resonance wavelength and impedance does not vary much.

Material	ϵ_r	ϵ_{reff}	Substrate Height H (cm)	Wavelength at Resonance, f=2GHz (cm)	Dipole Length (cm)	Zin Resonant (kohms)	Zin 2nd Harmonic (kohms)
Air	1	1	10	14.9	7.50	1.222	2.286
SiO_2	3.9	1	1	14.9	7.50	1.222	2.286
SiO_2	3.9	2.17	5	10.2	5.14	1.239	2.335
SiO_2	3.9	2.35	10	9.77	4.92	1.242	2.339
Alumina	9.4	1	1	14.9	7.50	1.222	2.286
Alumina	9.4	4.92	5	6.72	3.36	1.266	2.344
Alumina	9.4	5.2	10	6.53	3.26	1.268	2.340

While several antenna sizes are to be attempted on the wafer the first design is the baseline size to be built for the 2GHz/6GHz range. Given that the diode characteristics exhibit large third order mixing the antenna design focuses on the interrogation and second harmonic frequencies. This yields an ultimate length of the dipole of 5.14cm for resonance at 2GHz and 6GHz, assuming a 5cm glass base below the substrate. The thick base (sheets of glass) is desirable for testing since at higher frequencies the substrate plays an increasingly more important role in the operation of the antenna, wherein the effective resonant length is reduced. The thick base is a method of simulating the condition at lower frequencies. While the dipole length is very large for microfabrication purposes, it is a prototype system with the goal of further miniaturization by increasing the frequencies of operation. For example, for operation in the 24GHz amateur band the antenna size would be, at most, 1.25cm and only about 7mm if the effect of the substrate is taken into account. At that size, microfabrication is the most suitable method of creating the tags.

Alternatively, the tags designed by Dr. Chris Backhouse and Brian Moore had folded dipole lengths of just 3.1cm (and designed for 917MHz operation), which they had found to work well with an HP 2800 series diode at distances up to 12m from a Recco interrogation unit (917MHz/1.8GHz system). The 3.1cm dipole represents a dipole length of $\lambda/5$ at 917MHz, when the dielectric constant of the substrate (FR4, $\epsilon_r \approx 4.5$) is considered. The optimal dipole length of 3.1 cm was found empirically through tests in the field with the Recco system.

This approach should be feasible for use at 10GHz or even higher frequencies (even THz, though more stringent electromagnetic modeling may be desirable at these frequencies). An optimal resonant antenna could be used at these high frequencies, as the dimensions would make microfabrication of the antennas feasible. The length of the antenna would scale to smaller dimensions leading to a smaller signal available (due to a smaller effective length) to turn on the device (into the non-linear region). This effect could be countered with a higher gain antenna design or by biasing the device with a small on-chip battery into its non-linear region. Ultimately, it is desired to have a smaller linear region of the MIM diode and therefore a smaller turn-on voltage. This would enable ultra small transceiver devices that may be integrated with MEMS devices (i.e. smart dust).

9

Chapter 9: Conclusions and Future Work

9.1 Chapter Overview

This chapter summarizes the present work and suggests directions for future development

Nanoelectronic devices that are compatible with MEMS processes have been fabricated in a multi-user fabrication facility, the University of Alberta Nanofab. Devices with suitably non-linear characteristics were fabricated and a good yield was achieved for this type of device. The primary electrical problems associated with the manufacture of tunnel junctions were dielectric breakdown of devices and uniformity of *i-v* characteristics. The process itself is tolerant of low-resolution photolithography due to thick photoresist, low metal-substrate adhesion and high base pressures (for nanoelectronics) during metal deposition. For oxide thicknesses of 40Å an electron beam deposition rate of 1.5Å/s yielded devices with very non-linear characteristics and that were less susceptible to failure than 30Å oxide thickness devices or devices fabricated using higher deposition rates, due to fewer pin-hole sites as a result of a thicker and more uniform insulating layer. The devices functioned, in a non-linear *i-v* characteristic sense, at high voltages (>1V). The *i-v* characteristic data was fitted to Simmons' theory and good fits that were comparable to or better than other reports, were achieved. The F-N plots for these devices showed the expected characteristic behaviours. A starting point for future investigation into better quality diodes has been presented.

A fabrication process for a tunnelling transistor, the MITT, has been proposed based on the process developed for the MIM diode. The transistor follows closely the MIM diode cross-layer process and could be fabricated now using conventional photolithography and production capable equipment. Ultimately, with standard process refinement, this could lead to circuits of tunnelling devices fabricated using only microlithographic processes.

An R.F. tag application has also been designed for the diodes that can be entirely microfabricated on a Si substrate. The diode has been verified in operation up to 600MHz, limited by the maximum operating frequency of the testing equipment set-up. Second and fourth harmonic frequencies were observed in response to signals as small as –3dBm (50ohm system) and first harmonic frequencies were observed at input signals of 0dBm, verifying that the diode functions as a harmonic mixer and verifying its operation at R.F. frequencies.

9.2 Major Challenges

Several challenges arose during the research. First and foremost were challenges involving the suitability and availability of certain equipment for the fabrication and testing of the tunnelling devices. In a multi-user fabrication environment the equipment used easily becomes cross-contaminated with incompatible processes (or in the case of the e-beam system, by the pumping system itself). While this is difficult to prove conclusively, it was found to be a common problem amongst several users. For the purposes of tunnel junctions, with insulator thicknesses on the order of a few nanometres, any contamination could be critically detrimental to the proper functioning of the device. Ultimately, this issue was most responsible for the research not progressing further than desired and it will exist for the foreseeable future, though steps are being taken in the Nanofab to rectify them, such as installing equipment for process-specific applications or having 'clean' and 'dirty' systems classified according to the material for deposition (i.e. metals only vs. any material). Also, the adhesion issues mentioned in the fabrication section posed a problem for device fabrication. This issue has been ostensibly dealt with by changing the method of pumping from diffusion pumping (oil backstreaming possibly a cause of poor adhesion) to cryo-pumping. An issue that remains to be dealt with if the work is to be continued is the availability of R.F. test equipment, including appropriate probing. While it was possible to achieve some results at R.F. in this research, it will be necessary to have a superior set-up for future development as the future harmonic radar tag is to function at frequencies of 10GHz and greater.

9.3 Future Work

The current research can be extended with future work in several directions. First, the electrical characteristics of tunnel junctions could be studied as a function of different metals used for the electrodes. Not much information was found in the literature on this subject specifically (with most researchers choosing materials based on reasoned assumptions). As a result, a proven basis for material selection does not exist. Among the metals likely to be studied are nickel, aluminum, niobium, chromium and titanium, due to their oxidation characteristics. The tunnel junction quality could be examined in terms of breakdown voltage, non-linearity and ageing. The best quality junctions could be investigated as a function of materials used and oxidation parameters for the different metals. Furthermore, annealing could be investigated as a method to improve crystallinity of the junction oxide. A more crystalline oxide barrier would yield a more stable and reliable device, since charge trapping and electron scattering effects are minimized. It was found that even R.F. reactively sputter-deposited oxides can increase crystallinity by annealing in air at temperatures of 300-1100°C for just 1h.[115]

Other fabrication techniques, such as ion-beam etching could also be investigated. Ion beam etching has been used to fabricate step-edge SIS junctions[116], with steep sidewalls and room temperature MIM junctions as well.[117] Both the pattern etching and oxide growth could be done using ion-beam methods. This could potentially have several advantages over the RIE sputter-etch method presented in this research. First, steep sidewalls with no overlap of passivation layers are an intrinsic property of the process. Diodes fabricated entirely with photolithographic methods could have the smallest junction areas possible with this method. Second, with ion-beam oxidation, the oxide film has a higher probability of being a cleaner film since the sample is physically separated from the plasma and, unlike the RIE method, sputtering of contaminants from the chamber sidewalls is not an issue.

An MIM diode could be fabricated using ion-beam etching by depositing the electrode metal and a passivation SiO_2 layer using sputtering or e-beam evaporation onto an unpatterned SiO_2 covered wafer. A photoresist pattern can then be applied onto the deposited material and hard-baked. It is known that photoresist has a lower etch rate than either SiO_2 or metals so it can be used to pattern these materials.[118] After hard baking, the photoresist-covered wafer can be placed into an ion mill and exposed to an Ar^+ plasma, which etches the SiO_2/metal layer at a faster rate than the photoresist. This would create an edge MOM with a steep sidewall, and therefore a small junction area, suitable for high frequency applications. Alternatively, the nickel layer could be deposited first, and an SiO_2 passivation layer could then be patterned by lift-off in a subsequent step. The thick SiO_2 layer could then be used to mask the nickel in the ion beam etcher, assuming the nickel layer is much thinner. Without the photoresist during ion beam etching this could result in a cleaner nickel surface. Either of these approaches could be used to solve the lift-off problems of this research, as the adhesion would presumably become less of an issue. Oxidation of the metal using an ion-beam etcher simply involves exposing the sample to an oxygen ion beam. Metal oxide films are grown on the metal surface with their quality and thickness dependent on ion energy and exposure time.[117] This type of process can be readily implemented in the University of Alberta Nanofab.

Further investigation can be done into a more complex R.F. harmonic radar identification system. This could consist of a complete remote sensing tag with integrated MEMS and electronics. The asymmetric MIM diodes could be used as rectification in a ‘rectenna’ to power on-chip electronics while the symmetric diodes could be used to generate a response harmonic to an interrogation frequency signal. The tunnelling transistors and magnetic tunnel junctions could be investigated for integration into a system, such as a harmonic radar tag, which seems like an ideal application due to the low power and non-volatile nature of the MITT and MTJ components. This chip could integrate the nanoelectronics and MEMS with the harmonic radar components for a single chip solution. This would be advantageous over CMOS in that the nanoelectronic devices presented here are fabricated with a MEMS compatible process so that no flip-chipping would be required. Alternatively, with diodes that have a small enough junction area (and therefore small capacitance and higher maximum frequency of operation) efforts could be made to directly convert solar energy to D.C. voltage, as is already the goal for some research groups.[32, 45]

Finally, to properly advance in these directions it would be necessary to solve the remaining problems in the MIM diode fabrication; namely the lift-off lithography (fine features pattern by trying different photoresists and anti-reflection coatings and adhesion of metals to substrate and other metal by testing different thin (few nm) adhesion layers) and uniformity of device characteristics. This would entail more detailed development and characterization of the lithography as well as characterization of the deposition tools, possibly using materials science tools. Further investigation of the tunnel diode structure may be necessary to determine the quality of the tunnel junctions using Auger Electron Spectroscopy (AES), Electron Dispersive X-Ray Analysis (EDX) or Secondary Ion Mass Spectroscopy (SIMS) thin film material characterization equipment to determine the chemical composition of the tunnel oxide barrier. These methods would give a clear indication of what effect various process parameters have on junction quality.

9.4 Summary

This thesis presents the results of a MEMS compatible process for MIM diodes. Diodes with good characteristics were fabricated and were described well by quantum theoretical models. D.C. tests show a variation in *i-v* characteristics between devices, while R.F. tests verify that the device is a viable harmonic generator at R.F. frequencies. Two applications for the MIM diodes were presented; a harmonic radar tag was designed, based on the D.C. and R.F. testing of the MIM diodes, and a tunnelling transistor fabrication process was specified, based on the process for the MIM diodes.

Bibliography

---

[1] "International Technology Roadmap for Semiconductors," ITRS, 2002, p.48, available online at http://public.itrs.net.

[2] L. Chang and C. Hu, "MOSFET Scaling into the 10nm Regime," *Superlattices and Microstructures,* vol. 28, no. 5/6, pp. 351-355 (2000).

[3] D.J. Frank, R.H. Dennard, E. Nowak, P.M. Solomon, Y. Taur, H-S.P. Wong "Device Scaling Limits of Si MOSFETs and their Application Dependencies," *Proceedings of the IEEE*, vol. 89, no. 3, pp. 259-287 (2001).

[4] Y. Taur, D.A. Buchanan, W. Chen, D.J. Frank, K.E. Ismail, S-H Lo, G.A. Sai-Halasz, R.G. Viswanathan, H-J.C. Wann, S.J. Wind, H-S. Wong "CMOS Scaling into the Nanometer Regime," *Proceedings of the IEEE,* vol. 85, no. 4, pp. 486-504 (1997).

[5] R. Doering, Y. Nishi, "Limits of Integrated Circuit Manufacturing," *Proceedings of the IEEE,* vol. 89, no. 3, pp. 375-393 (2001).

[6] B.N. Engel, N.D.Rizzo, J. Janesky, J.M. Slaughter, R. Dave, M. DeHerrera, M. Durlam, "The Science and Technology of Magnetoresistive Tunneling Memory," *IEEE Transactions on Nanotechnology,* vol. 1, no. 1, pp. 2-38 (2002).

[7] I Wilke, Y Oppliger, W. Herrmann, and F.K. Kneubühl, "Nanometer thin-film Ni-NiO-Ni diodes for detection and mixing of 30THz radiation," *Infrared Physics and Technology,* vol. 39 (3) pp. 123-183 (Apr 1998).

[8] A. Hoofring and V. Kapoor, "Submicron Nickel-Oxide-Gold Tunnel Diode Detectors for Rectennas," *Journal of Applied Physics*, vol. 66 (11) pp. 430-437 (1989).

[9] J. Simmons, "Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film," *Journal of Applied Physics,* vol..34 (6) pp. 1793-1803 (1963).

[10] N. Zettili, "Quantum Mechanics, Concepts and Applications," John Wiley & Sons Ltd, Chichester, UK, 2001.

[11] L.I. Schiff, "Quantum Mechanics, Second Edition," McGraw-Hill, New York, 1955.

[12] W.V. Houston, G.C. Phillips, "Principles of Quantum Mechanics," North-Holland Publishing Company, Amsterdam, 1973.

[13] D.S. Saxon, "Elementary Quantum Mechanics," Holden-Day, San Francisco, 1968.

---

[14] B. Streetman, “Solid-State Electronic Devices, Fourth Edition” Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

[15] “Tunneling Phenomena in Solids,” E. Burstein, S. Lundqvist, ed., p.p. 35, Plenum Press, New York, 1969.

[16] J.C. Fisher, I. Gaiever, “Tunneling Through Thin Insulating Layers,” *Journal of Applied Physics,* vol. 32. no. 2, pp. 172-177 (1961).

[17] J. Simmons, “Electric Tunnel Effect Between Dissimilar Electrodes Separated by a Thin Insulating Film,” *Journal of Applied Physics,* vol. 34 (9) pp. 2581-2590 (1963).

[18] J. Simmons, “Generalized Thermal J-V Characteristic for the Electric Tunnel Effect,” *Journal of Applied Physics*, vol. 35 (9) pp. 2655-2658 (1964).

[19] R.J. Soukup, “Schottky Emission in Thin Film Sandwich Structures,” *Journal of Applied Physics,* vol. 46, no. 1, pp. 463-464 (1974).

[20] W.F. Brinkman, R.C. Dynes, J.M. Rowell, “Tunneling Conductance of Asymmetrical Barriers,” *Journal of Applied Physics,* vol. 41, no. 5, pp. 1915-1921 (1970).

[21] Z. Hurych, “Influence of Non-Uniform Thickness of Dielectric Layers on Capacitance and Tunnel Currents,” *Solid-State Electronics,* vol. 9, pp. 967-979 (1966).

[22] L.S. Dorneles, D.M. Schaefer, M. Carara, L.F. Schelp, ”The Use of Simmons’ Equation to Quantify the Insulating Barrier Parameters in Al/AlOx/Al Tunnel Junctions,” *Applied Physics Letters,* vol. 82 (17), pp. 2832-2834 (2003).

[23] H.Hübers, G.W. Schwaab, H.P. Röser, “Video Detection and Mixing Performance of GaAs Schottky-Barrier Diodes at 30THz and Comparison with Metal-Insulator-Metal Diodes,” *Applied Physics Letters,* vol. 75, no. 8, pp. 4243-4248 (1994).

[24] B.L. Twu, S.E. Schwarz, “Mechanism and Properties of Point Contact Metal-Insulator-Metal Diode Detectors at 10.6μ,” *Applied Physics Letters,* vol. 25, no. 10, pp. 595-598 (1974).

[25] Y.Yasuoka, T.Sakurada, T.Miyata, T.K. Gustafson, “Comparison of the Responsivity of W-Ni Point Contact MBM Diodes with W-Si Point Contact Schottky Diodes,” *IEEE Journal of Quantum Electronics,* vol. QE-16, no. 12, pp. 1372-1375 (1980).

[26] R.H. Havemann, “Photolithographic fabrication of Thin-Film Metal-Oxide-Metal Diodes with Submicrometer-square Junction Areas,” *Journal of Vacuum Science and Technology,* vol.15 (2), pp. 389-391 (1978).

---

[27] S.K. Masalmeh, H.K.E. Stadermann, J. Korving, "Mixing and Rectification Properties of MIM Diodes," *Physica B*, vol. 218, pp. 56-59, (1996).

[28] M. Heiblum, S Wang, J Whinnery, and T. K. Gustafson, "Characteristics of Integrated MOM Junctions at dc and at Optical Frequencies," *IEEE Journal of Quantum Electronics,* vol.14 (3), pp. 159-169 (March 1978).

[29] H.D. Riccius, "High Frequency Limitation of Metal-Insulator-Metal Point Contact Diodes," *Applied Physics Letters,* vol .27, no. 4, pp. 232-233 (1975).

[30] T. Gustafson, R.V. Schmidt, J.R. Perucca, "Optical Detection in Thin-Film Metal-Oxide-Metal Diodes," *Applied Physics Letters,* vol. 24 (12), pp. 620-621 (1974).

[31] H. Scherer, T. Weimann, P. Hinze, B.W. Sawyer, A.B. Zorin, J. Niemeyer, " Characterization of all chromium tunnel junctions and single-electron tunneling devices by direct-writing multilayer technique" *Journal of Applied Physics,* vol. 86 (12) pp. 6956-6964 (1999).

[32] E. Grossman, T.E. Harvey, C.D. Reintsema, "Controlled Barrier Modification in Nb/NbOx/Ag Metal Insulator Metal Tunnel Diodes," *Journal of Applied Physics,* vol. 91 (12), pp. 10134-19139 (2002).

[33] S. Morohashi, F. Shinoki, A. Shijo, M. Aoyagi, H. Hayakawa, "High Quality Nb/Al-AlOx/Nb Josephson Junction," *Applied Physics Letters,* vol. 46. no. 12, pp. 1179-1181 (1985).

[34] M. Aoyagi, M. Maezawa, H. Nakagawa, I. Kurosawa, "Fabrication of Submicron Nb/AlOx/Nb Josephson Junctions Using ECR Plasma Etching Technique," *IEEE Transactions on Applied Superconductivity,* vol. 7, no. 2, pp. 2644-2648 (1997).

[35] V. Patel, J.E. Lukens, "Self-Shunted Nb/AlOx/Nb Josephson Junctions," *IEEE Transactions on Applied Superconductivity,* vol. 9, no. 2, pp. 3247-3250 (1999).

[36] M.P. Janawadkar, R. Baskaran, K. Gireesan, R. Saha, L.S. Vaidyanathan, T.S. Radhakrishnan, "High Quality Nb/Al-AlOx-Al/Nb Josephson Junctions by Electron Beam Evaporation," *Japanese Journal of Applied Physics Part 2: Letters,* vol. 33. no. 12A, pp. L1662-L1664 (1994).

[37] J.R. Tucker, C. Wang, P.S. Carney, "Silicon Field-Effect Transistor Based on Quantum Tunneling," *Applied Physics Letters*, vol. 65, no. 5, pp. 618-620 (1994).

[38] K. Fujimaru and H. Matsumura, "Theoretical Consideration of a New Nanometer Transistor Using Metal/Insulator Tunnel-Junction," *Japanese Journal of Applied Physics,* vol. 35 (1) no. 4A, pp. 2090-2094 (1996).

---

[39] E.S. Snow, P.M. Campbell, R.W. Rendell, F.A. Buot, D. Park, C.R.K. Marrian, R. Mango, "A Metal/Oxide Tunneling Transistor," *Applied Physics Letters,* vol. 72 (23), pp. 3071-3073 (1998).

[40] K. Fujimaru, R. Sasajima, H. Matsumura, "Nanoscale Metal Transistor Control of Fowler-Nordheim Tunneling Currents Through 16 nm Insulating Channel," *Journal of Applied Physics,* vol. 85 (9), pp. 6912-6916 (1999).

[41] K. Fukushima, R. Sasajima, K. Fujimaru, H. Matsumura, "A Novel Nanoscale Metal Transistor Fabricated by Conventional Photolithography," *Japanese Journal of Applied Physics,* vol. 38, no. 12B, pp. 7233-7236 (1999).

[42] J.G. Roland, G. McKinnon, C. Backhouse, P.D. Taylor, "Even Smaller Radar Tags on Insects," *Nature,* vol. 381, pp. 120 (1996).

[43] D.Landis, E. Rothwell, D. Reinhard, L. Kempel, B. Wilmhoff, M. O'Neal, J. Miller, "Development of a harmonic radar system for tracking beneficial insects between refuge and crop habitats," Landscape Ecology and Biology Lab Research Abstracts, Michigan State University Internet (World Wide Web), 2001, http://www.cips.msu.edu/landislab/abstracts.html

[44] L. Perrier, "Le Radar Harmonique Comme Technique de Pistage des Petits Animaux," Bibliographic Research Report, Ecole Nationale Supérieur Des Sciences de l'Information et des Bibliothequès, Université Claude Bernard, Lyon (2001).

[45] B. Berland *et al.*, "Optical Rectenna for Direct Conversion of Sunlight to Electricity," Proceedings of the National Center for Photovoltaics Program Review Meeting, p. 323 (2001)

[46] C. Fumeaux, J. Alda, G.D. Boreman, "Lithographic Antennas at Visible Frequencies," *Optics Letters*, vol. 24, no. 22, pp. 1629-1631 (1999).

[47] R. Corkish, M.A. Green, T. Puzzer, "Solar Energy Collection by Antennas," *Solar Energy,* vol. 73 (6), pp. 395-401 (2002).

[48] A. Marks, "Femto Diode and Applications," U.S. Patent #4,720,642 (1988).

[49] B. Berland, "Photovoltaic Technologies Beyond the Horizon," Report on commissioned research for the NREL (Report NREL/SR-520-33263) (2003). Also available at http://www.nrel.gov/docs/fy03osti/33263.pdf.

[50] S. Mickan, D. Abbott, J. Munch, X-C. Zhang, T. van Doorn, "Analysis of System Trade-Offs for Terahertz Imaging," *Microelectronics Journal,* vol. 31, pp. 503-514 (2000).

---

[51] C. Zandonella, “T-Ray Specs,” *Nature,* vol. 434 (6950), pp. 721-722 (2003).

[52] X-C. Zhang, “Terahertz Wave Imaging: Horizons and Hurdles,” *Physics in Medicine and Biology,* vol. 47,no. 7, pp. 3667-3677 (2002).

[53] G. Zhao, R.N Schouten, N. van der Valk, W.T. Wenckebach, P.C.M. Planken, “A Terahertz System Using Semi-Large Emitters: Noise and Performance Characteristics,” *Physics in Medicine and Biology*, vol. 47, no. 7, pp. 3699-3704 (2002).

[54] W.J. Gallagher, S.S.P. Parkin, Y. Lu, X.P. Bian, A. Marley, K.P. Roche, R.A. Altman, S.A. Rishton, C. Jahnes, T.M. Shaw, G. Xiao, “Microstructured Magnetic Tunnel Junctions (invited),” *Journal of Applied Physics,* vol. 81 no. 8, pp. 3741-3746 (1997).

[55] S. Tehrani B. Engel, J. M. Slaughter, E. Chen, M. DeHerrera, M. Durlam, P. Naji, R. Whig, J. Janesky, J. Calder, “Recent Developments in Magnetic Tunnel Junction MRAM,” *IEEE Transactions on Magnetics,* vol. 36, no. 5, pp. 2752-2757 (2000).

[56] J.C. Gallop, “SQUIDS, the Josephson Effect and Superconducting Electronics,” Adam Hilger, Philadelphia, 1991.

[57] C.A. Hamilton, C.J. Burroughs, S.P. Benz, “Josephson Voltage Standard – A Review,” *IEEE Transactions on Applied Superconductivity,* vol. 7. no. 2, pp. 3756-3761 (1997).

[58] T. Nagatsuma, K. Enpuku, F. Irie, K. Yoshida, “Flux-Flow Type Josephson Oscillator for Millimeter and Submillimeter Wave Region. IV. Thin-Film Coupling” *Journal of Applied Physics,* vol. 54. no. 6, pp. 3302-3309 (1983).

[59] V. Kaplunenko, “On Chip Clock Oscillator for High Precision RSFQ Applications,” *IEEE Transactions on Applied Superconductivity,* vol. 13, no. 2, pp. 575-578 (2003).

[60] S.T. Ruggerio,D.A. Rudman, ed. “Superconducting Devices,” Academic Press, Boston, 1990.

[61] A.C. Callegari, R.A. Buhrman, “Millimeter Wave Mixing with Submicron Area Nb Tunnel Junctions,” *Journal of Applied Physics,* vol. 53, no.2, pp. 823-827 (1982).

[62] B.D. Jackson, T.M. Klapwijk, “The Current Status of Low-Noise THz Mixers Based on SIS Junctions,” *Physica C,* vols. 372-376, pp. 368-373 (2002).

[63] S. Haraichi, T. Wada, S.M. Gorwadkar, K. Ishii, “Fabrication of Nanometer-Scale Vertical Metal-Insulator-Metal Tunnel Junctions Using a Silicon-on-Insulator Substrate,” *Japanese Journal of Applied Physics,* vol. 37, no. 3B, pp. 1580-1583 (1998).

---

[64] C. Oviedo, "Oxidation Kinetics of Pure Titanium at Low Pressures," *Journal of Physics: Condensed Matter,* vol. 5, pp. A153-A154 (1993).

[65] P.C. Karulkar, J.E. Nordman, "Fabrication of Nb-NbOx-Pb Josephson Tunnel Junctions using RF Glow Discharge Oxidation," *Journal of Applied Physics,* vol. 50, no. 11, pp. 7051- 7059 (1979).

[66] S.V. Kumari, M. Natarajan, V.K. Vaidyan, P. Koshy, "Surface Oxidation of Nickel Thin Films," *Journal of Materials Science Letters,* vol. 11, no. 11, pp. 761-762 (1992).

[67] D.R. Baraff, R. Boynton, B. Gribbon, J.R. Long, B.K. MacLaurin, C.J. Miner, N.M. Serinken, R.W. Streater, W.D. Westwood, "Application of Metal-Insulator-Metal Nonlinear Devices in Multiplexed Liquid Crystal Displays," *Society for Information Display International Symposium Digest of Technical Papers, Florida, 1980 (Society for Information Display, Playa del Rey, 1980) p200.*

[68] S. Zhang, J.K.O. Sin, "A Novel Self-Aligned Bidirectional MIM Diode with Transparent Junctions for AM-LCD's," *IEEE Electron Device Letters,* vol. 19, no. 6, pp. 192-194 (1998).

[69] V. Hochholzer, E. Leuder, T. Kallfass, "A Two-Mask Fabrication of Ta_2O_5 MIMs with Adjustable Threshold and Gain," *SID 90 Digest,* pp. 526-529 (1990).

[70] T. Hirai, H. Morita, M. Tasaka, "Current-Voltage Characteristics and Conductivity of Metal-Insulator-Metal (Ti/TaOx/Ta) Type Thin-Film Diode," *Japanese Journal of Applied Physics,* vol. 39, pp. 657-661 (2000).

[71] H. Liu, Y. Wu, J. Yuan, K. Ma, X. Huang, "Effect of Utilizing Hydrogen-Treated Tantalum Anodized Oxidation on Symmetry of Current-Voltage Characteristic of Metal-Insulator-Metal Element," *Japanese Journal of Applied Physics,* vol. 36, pp. 7334-7336 (1997).

[72] R.B. Laibowitz, J.J. Cuomo, "Tunneling Sandwich Structures Using Single-Crystal Niobium Films," *Journal of Applied Physics,* vol. 41, pp. 2748-2750 (1970).

[73] G. Hawkins, J. Clarke, "Nb-Nb Thin Film Josephson Junctions," *Journal of Applied Physics,* vol. 47, no. 4, pp. 1616-1619 (1975).

[74] T. Noguchi, H. Maezawa, S-C. Shi, S. Yamamoto, "Nb-Based Mixer in the 800 GHz Band," *IEEE Transactions on Applied Superconductivity,* vol. 11, no. 1, pp. 167-170 (2001).

[75] R.F. Broom, R.B. Laibowitz, Th.O. Mohr, W. Walter, "Fabrication and Properties of Niobium Josephson Tunnel Junctions," *IBM Journal of Research and Development,* vol. 24, no. 2, pp. 212-234 (1980).

---

[76] M. Gurvitch, M.A. Washington, H.A. Huggins, "High Quality Refractory Josephson Tunnel Junctions Utilizing Thin Aluminum Layers," *Applied Physics Letters,* vol. 42, no. 5, pp. 472-474 (1983).

[77] K.L. Westra, B.G. Veidt, J.F. Vaneldik, D. Routledge, M.J. Brett, "Fabrication of NbN/MgO/NbN Tunnel Junctions," *Physica C: Superconductivity,* vol. 181, no. 1-3, pp. 30-36 (1991).

[78] B.G. Veidt, "Enhancements to Millimetre- and Submillimetre-Wave Integrated Circuit Receivers," Ph.D. Thesis, University of Alberta, 1995.

[79] Z.H. Barber, M.G. Blamire, R.E. Somekh, J.E. Evetts, "Fabrication and Characterization of All-Refractory NbCN/Al/AlOx/Al/Nb Junctions," *IEEE Transactions on Applied Superconductivity,* vol. 3, no. 1, pp. 2054-2057 (1993).

[80] Heidelberg Instruments, Heidelberg, Germany.

[81] Tanner Research, Pasadena, CA, USA.

[82] Filmetrics Inc. San Diego, CA, USA.

[83] Arch Chemicals, Norwalk, CT, USA.

[84] Microchem Corp., Newton, MA, USA.

[85] OCG Microelectronic Materials, Tempe, AZ, USA.

[86] K.J. Lesker Co. Pittsburgh, PA, USA.

[87] "Handbook of Chemistry and Physics, 84th edition" Chemical Rubber Company, 2003.

[88] F.M. Smits, "Measurement of Sheet Resistivities with the Four-Point Probe," *Bell System Technical Journal*, vol. 37, pp. 711-718, 1958.

[89] A. Misra, M.F. Hundley., D. Hristova, H. Kung, T.E. Mitchell, M. Nastasi, J.D. Embury, "Electrical resistivity of sputtered Cu/Cr multilayered thin films," *Journal of Applied Physics,* vol. 85, no. 1, pp. 302-309 (1999).

[90] M. Hatzakis, B.J. Canavello, J.M. Shaw, "Single-Step Optical Lift-Off Process," *IBM Journal of Research and Development,* vol. 24 no. 4, pp. 452-460 (1980).

[91] R.M.A. Azzam, N.M. Bashara, "Ellipsometry and Polarized Light," North Holland Publishing Company, Amsterdam, 1977.

---

[92] A. Rothen, "Measurement of the Thickness of Thin Films by Optical Means, From Rayleigh and Drude to Langmuir, and the Development of the Present Ellipsometer," in "Ellipsometry in the Measurement of Surfaces and Thin Films," E. Passaglia, R.R. Stromberg and J. Kruger ed. Symposium Proceedings, National Bureau of Standards (US) Washington, 1963.

[93] Geartner Model L125B, Gaertner Scientific, Chicago, IL, USA

[94] J.A. Woollam, Lincoln, NE, USA

[95] L.M.M. de Souza, F.P. Kong, F.R. McLarnon, R.H. Muller, "Spectroscopic Ellipsometry Study of Nickel Oxidation in Alkaline Solution," *Electrochimica Acta*, vol. 42 (8), pp. 1253-1267 (1997).

[96] L. Hanekamp *et al.* "Spectroscopic Ellipsometric Investigation of the clean and Oxygen Exposed Ni(110) Surface," *Surface Science*, vol. 135, pp. 383-395 (1983).

[97] Y. Ma, T. Tsurumi, S. Nishizawa, N. Ohashi, O. Fukunga, "Spectroscopic Ellipsometry of Nickel Oxide/Zinc OxideArtificial Superlattices," *Journal of the American Ceramic Society*, vol. 81 (8), pp. 2125-2129 (1998).

[98] J.S. Hale, M. DeVries, B. Dworak, J.A. Woollam, "Visible and Infrared Optical Constants of Electrochromic Materials for Emissivity Modulation Applications," *Thin Solid Films*, vol. 313-314, pp. 205-209 (1998).

[99] M. Wada, J. Nakano, F. Yanagawa, "Inhibiting Effect of Thermal Oxides on the Growth of RF-Plasma Oxides for Pb Alloy Josephson Junctions," *Journal of Applied Physics*, vol. 54, no. 11, pp. 6658-6662 (1983).

[100] N. Cabrera, N.F. Mott, "Theory of the Oxidation of Metals," *Reports on Progress in Physics*, vol. 12, pp. 163-184 (1948).

[101] J. Greiner, "Josephson Tunneling Barriers by rf Sputter Etching in an Oxygen Plasma," *Journal of Applied Physics*, vol. 42 (12), pp. 5151-5155 (1971).

[102] Private communication with Dr. Doug Vick, Edmonton, AB.

[103] M.J. Madou, "Fundamentals of Microfabrication," CRC Press, 1997.

[104] G. Wakefield, P.J.Dobson, Y.Y. Foo, A. Loni, A. Simons, J.L. Hutchison, "The Fabrication and Characterization of Nickel Oxide Films and their Application as Contacts to Polymer/Porous Silicon Electroluminescent Devices," *Semiconductor Science and Technology*, vol. 12, pp. 1304-1309 (1997).

---

[105] J.D.R. Buchanan, T.P.A. Hase, B.K. Tanner, "Determination of the Thickness of Al_2O_3 Barriers in Magnetic Tunnel Junctions," *Applied Physics Letters,* vol. 81, no. 4, pp. 751-753 (2002).

[106] E.Z. Luo, S.K. Wong, A.B. Pakhomov, J.B. Xu, I.H. Wilson, C.Y. Wong, "Tunneling Current and Thickness Inhomogeneities of Ultra-Thin Aluminum Oxide Films in Magnetic Tunnel Junctions," *Journal of Applied Physics,* vol. 90, no. 10, pp. 5202-5207 (2001).

[107] ANSYS Inc, Canonsburg, PA, USA

[108] Private communication with Brian Moore, Edmonton, AB.

[109] J.D. Kraus, R.J. Marhefka, "Antennas for All Applications: Third Edition" McGraw-Hill, pp. 594, 2002.

[110] A.W. Rudge, K.Milne, A.D. Olver, P. Knight, *ed.* "The Handbook of Antenna Design: Volume 2," Peter Peregrinus Ltd., London, 1983.

[111] C. Balanis, "Antenna Theory," Harper and Row, New York, 1982.

[112] A. Sanchez, C.F. Davis Jr., K.C. Liu, A. Javan, "The MOM Tunneling Diode: Theoretical Estimate of Its Performance at Microwave and Infrared Frequencies," *Journal of Applied Physics,* vol. 49, no. 10, pp. 5270- 5277 (1978).

[113] "The Zero Bias Schottky Detector Diode," H.P. Application Note 969, Hewlett-Packard.

[114] R. Garg, P. Bhartia, I. Bahl, A. Ittipiboon, "Microstrip Antenna Design Handbook," Artech House, Boston, 2001.

[115] N. Martin, C. Rousselot, D. Rondot, F. Palmino, R. Mercier, "Microstructure Modification of Amorphous Titanium Oxide Thin Films During Annealing Treatment," *Thin Solid Films,* vol. 300, no.1-2, pp. 113-121 (1997).

[116] J.B. Bulman, O. Salazar, "AFM Analysis of Step-Edge Josephson Junctions," *IEEE Transactions on Applied Superconductivity,* vol. 7 no. 2, pp. 3650-3653 (1997).

[117] J.M.E. Harper, M. Heiblum, J.L.Speidell, J.J. Cuomo, "Ion Beam Oxidation," *Journal of Applied Physics,* vol. 52, no. 6, pp. 4118-4121 (1981).

[118] Private communication with University of Alberta Nanofab staff.

Appendix A: Equipment and Set-Up Specifications

RIE: Plasmatherm µEtch RIE
RF Source: ENI ACG-3XL (350W max), 13.56MHz through aluminum chuck
Gas inputs 1: CHF_3
2: O_2
3: CF_4
4: Ar/SF_6 selectable by T switch
General purpose etcher used mainly for oxide etches using CF_4 and SF_6.
Rough pumping by Leybold D40BCS (32.5cfm)
High Vac pumping by Leybold RUVAC WS 251 Roots Blower (164cfm)

Sputtering "Doug":
Kurt J. Lesker Inc. CMS-18: 2 target sputtering system (targets can be separated by a cross-contamination shield)
Target-Sample distance ~ 10cm.
Pumped by an APD Cryogenics Marathon® cryopump.

E-Beam Evaporation "Gomez": In-House Constructed
10" NRC model# 0178 Diffusion pump
8MHz Quartz crystal thickness monitor

Ellipsometer: Gaertner L-125B Laser Ellipsometer (Rotating Analyzer)
$\lambda = 632.8$nm

Spectroscopic Ellipsometer: J.A. Woolham VASE

Mask Aligners:

"Elmo": ABM Contact Mode Mask Aligner with Karl Suss Optical microscope alignment
UV exposure intensity at 365 and 400nm (used simultaneously):
365nm: 27mW/cm^2
400nm: 60.3mW/cm^2
"Oscar": ABM Contact Mode Mask Aligner with CCD camera alignment
UV exposure intensity at 365 and 400nm (used simultaneously):
365nm: 20.5mW/cm^2
400nm: 59mW/cm^2

Semiconductor Parameter Analyzer: HP 4145A
4 programmable current and voltage source/measurement units (SMU1-SMU4)
2 programmable voltage sources (VS1, VS2)
2 programmable voltage monitors (VM1, VM2)
HP "LIF" format 5 ¼" floppy disk drive

Network Analyzer: HP 8702B Lightwave Component Analyzer (in E/E mode)
300kHz-3GHz
50Ω input/output resistance, N-type connector
Spectrum Analyzer: Sony Tektronix 3086, 0-3GHz,
50Ω input resistance, N-type connector

Directional Coupler: Mini-Circuits ZEDC-10-2B
1-1000MHz
11 dB coupling +/- 0.5dB
18-35dB directivity
1.5W max input

Wafer Probes: Cascade Microtech – Alessi MD Series micropositioners with LEMO to BNC connector cabling (6").

RF cabling: 50Ω SMA male terminated cable, 1' and 2' lengths

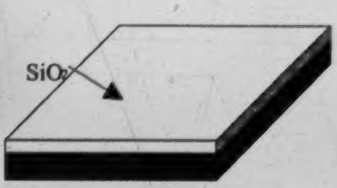
Appendix B: Step-Edge Junction Fabrication Process

For the higher frequency/smaller junction area edge-junction diodes it is necessary to use e-beam evaporation to deposit at least some of the layers. This is because film edges of sputtered films taper more slowly (have less steep transitions) in comparison to e-beam evaporated edges. This presents a problem for sputtered films in the edge diode configuration for the passivation SiO_2 layer on top of the bottom electrode since the possibility that some of the SiO_2 is deposited along the edge region increases. By using e-beam evaporation for the SiO_2 deposition step it is virtually assured that the electrode edges remain free from deposited SiO_2. The photomasks (Appendix C) were designed such that either the cross-layer process or this step-edge process can be fabricated using the same set of masks.

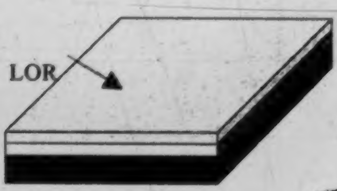
Process Steps:

1

2



3

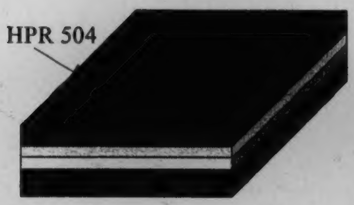


Note: Process steps 1-5 are the same as in the Cross-Layer process.

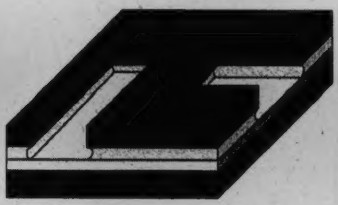
1. A bare silicon wafer is first soaked in hot Piranha ($3H_2SO_4{:}1H_2O_2$) for 30 minutes to remove contamination. It is then rinsed with de-ionized (D.I.) water in drop-deck and dry in spin-rinse dryer.

2. An insulating substrate is required for the devices. Rather than use expensive sapphire substrate or relatively rough glass, SiO_2 is grown on a silicon wafer through wet thermal oxidation. The wafers are placed in a furnace with a nitrogen bubbler for two intervals, between which they are turned 180° and re-oriented front to back for better uniformity. Typical parameters are 2.5 hours intervals at 1000°C to yield a layer of ~800nm of SiO_2. The thickness is a non-critical parameter but to decrease capacitance thicker is better so the time in the furnace and temperature can be increased.
3. After a dehydration bake of 5-10 minutes on a 200ºC hotplate. The wafer is then coated with Microchem's Lift-off resist 5B (LOR 5B) (dispensed at 500 rpm for 5s and spun at 3000 rpm for 45s) and pre-baked on a hotplate at 150°C for 5 minutes.

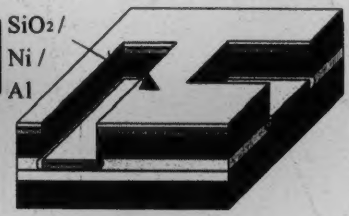
4



5



6

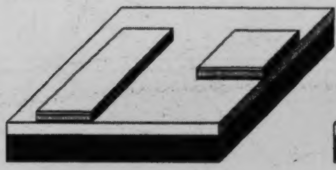


4. Photoresist (HPR 504) is spun at a rate of 5000 rpm to yield a photoresist thickness of 1.2μm. The photoresist is pre-baked on a hot-plate for 90 seconds at 110ºC and cooled for 1 minute at room temp.
5. The photoresist is exposed to mask 1 (large features) for 4s and developed with MF-354 developer. The wafer is then taken back to the vacuum hotplate and hard-baked at 120ºC for 2 minutes. The LOR is etched for an undercut with a

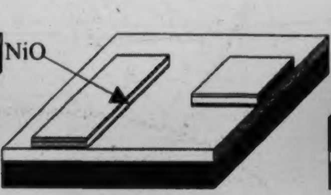
tetramethylammonium hydroxide (TMAH) based developer (MF-321) for 1 minute (yielding an undercut of approximately 5μm). Due to the hard-bake the HPR-504 photoresist should not be etched significantly by the TMAH based developer. The wafer is then flooded for 30s (UV exposed again without a mask to increase photoresist solubility for easier removal)

6. The wafer is taken to the electron beam evaporator where a thin layer (5-10nm) of aluminium is deposited as an adhesion layer, since nickel adheres poorly to silicon dioxide, followed by a layer of nickel (<150nm… expect better device performance with lower contact area, i.e. thinner nickel layer) without breaking vacuum. In each evaporation step the metals are melted with the shutter closed to the wafer line of sight until a clean evaporation melt surface is achieved (typically a couple of minutes at supply currents of 50A for Al and >100A for Ni). A final layer of SiO_2 is deposited to a thickness of ~50nm. The base pressure is $\sim 5\cdot 10^{-7}$ torr (before any of the three layers are deposited) and the nominal rate of deposition is 5Å/s for the Ni layer.

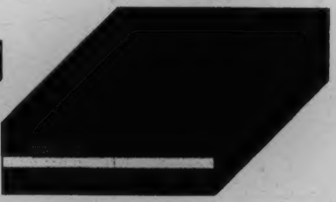
7



8



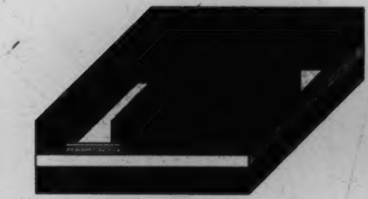
9



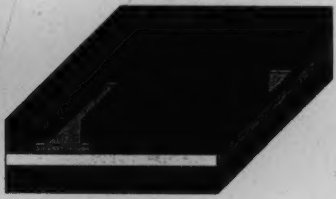
7. The photoresist is dissolved with a photoresist stripper for 10 minutes. The LOR resist is dissolved with LOR microstrip (Microchem Remover PG). The wafer is cleaned with DI water in the dump-rinse and spin rinse dryer.

8. The wafer is taken to a reactive ion etcher (RIE) that has been thoroughly cleaned with isopropyl alcohol (IPA) and a 100% O_2 200W plasma for a minimum of 30 minutes. The remnant O_2 plasma is evacuated and the chamber is preconditioned with 100% Ar plasma at 100W for 5 minutes. The wafer is then placed in the chamber and cleaned of native oxide with an Ar plasma at 100W for 10-15 minutes. A layer of NiO is grown by exposing the wafer to an Ar:O_2 ratio and power to yield a desired NiO thickness (i.e. 95:5 Ar:O_2 plasma at 45W and 150mtorr for a time exceeding 5 minutes yields a 30Å thick NiO layer)
9. HPR 504 Photoresist is spun on as in step 4.

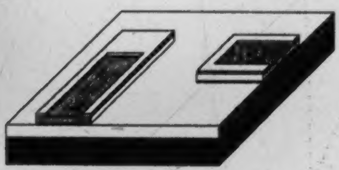
10



11

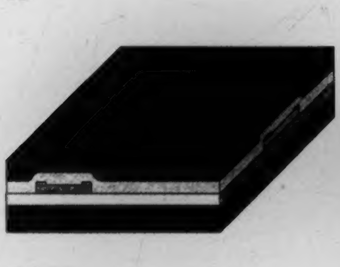


12

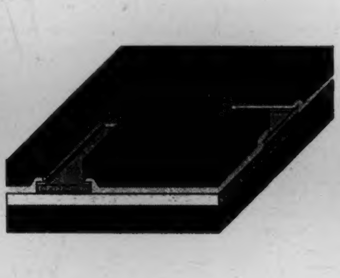


10. The photoresist (no LOR this time) is patterned with the pads mask (pad windows mask) and developed as before (or with the usual NaOH)
11. The wafer is then taken back to the RIE to selectively remove the evaporated silicon dioxide. Process parameters for this process are as follows: 100% tetrafluoromethane (CF_4) at 100W for less than 2 min (etched a 1.5µm hole in SiO_2 after 2 minutes with little effect to photoresist
12. Photoresist is stripped with acetone and the wafer is cleaned by rinsing in the dump-rinse and spin-rinse dryer.

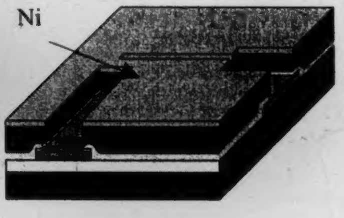
13



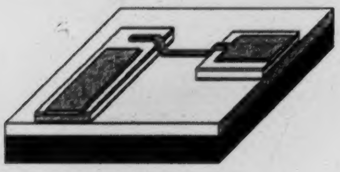
14



15

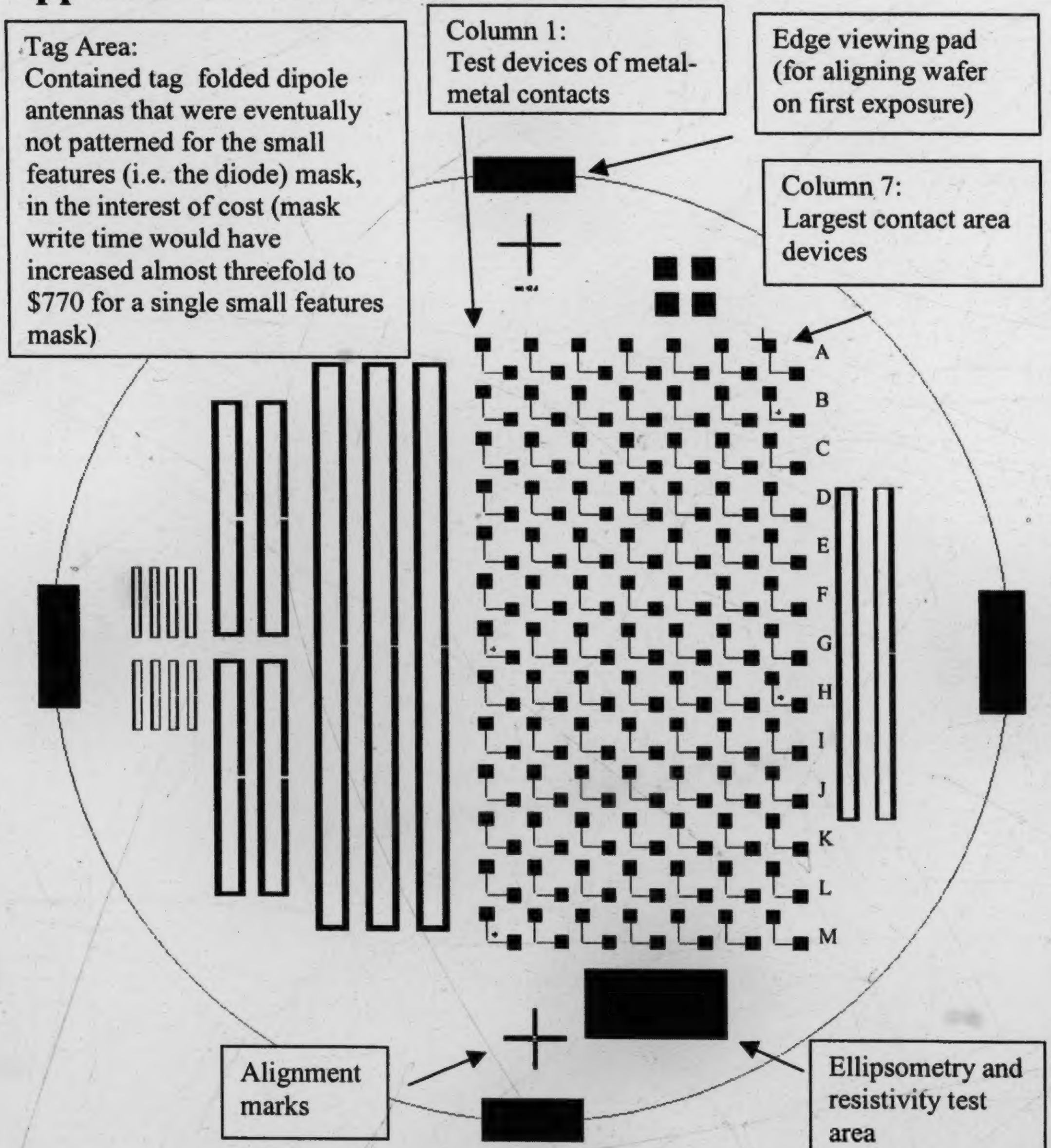


16



13. LOR 5B and HPR 504 photoresist are re-spun as in steps 3 and 4.
14. The wafer is exposed for 4 seconds each to the second electrode and pad windows masks (Masks exposed successively). The photoresist is developed in MF-354 for 15s and then hard-baked for 2 minutes at 120°C. The LOR is developed as before with MF- 321 for 1minute after the hard-bake.
15. Aluminum and nickel are sputtered onto the wafer to thicknesses of approximately 10nm and 200nm respectively. For better step coverage the deposition pressure can be higher than the 0.7mtorr. Initially lower power is used in the deposition until a sufficient thickness is achieved so as not to damage the thin NiO layer by ion bombardment.
16. The photoresist is removed as before. The wafer is cleaned in the dump-rinse and spin rinse dryer.

Appendix C: MIM Diode Process Masks



Tag Area:
Contained tag folded dipole antennas that were eventually not patterned for the small features (i.e. the diode) mask, in the interest of cost (mask write time would have increased almost threefold to $770 for a single small features mask)

Column 1:
Test devices of metal-metal contacts

Edge viewing pad (for aligning wafer on first exposure)

Column 7:
Largest contact area devices

Alignment marks

Ellipsometry and resistivity test area

Figure A - 1 MIMSLF Mask

MIMSLF – Large features mask. All pads, Test area pads and wafer edge-centering window. This mask represents the first mask in the deposition sequence, such that the features shown here ultimately result in the bottom layer of deposited metal. All features on this mask were written with a 40mm lens and are at least 10μm in dimension. The mask includes all features of the device except for the top layer electrode (which has a minimum feature size of 1μm).

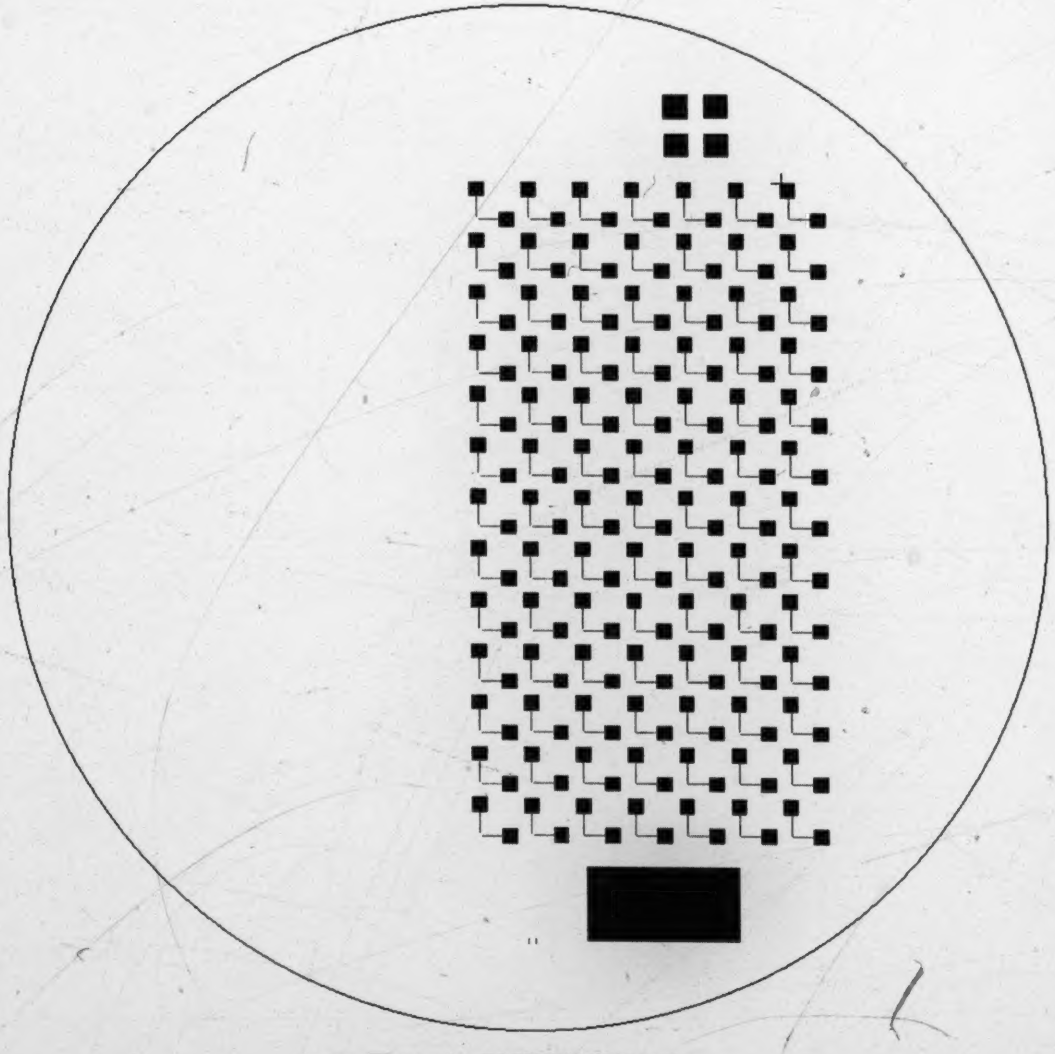


Figure A - 2 MIMSBOE Mask

MIMSBOE – Etch back mask for high-frequency process, for vias in oxide and to define areas where a final deposition of metal can serve to thicken pads. Each feature ultimately defines an etch window in the photoresist so that the oxide present on the device (plasma oxide in the cross-layer or SiO_2 in the edge junction) can be removed in a subsequent etch step.

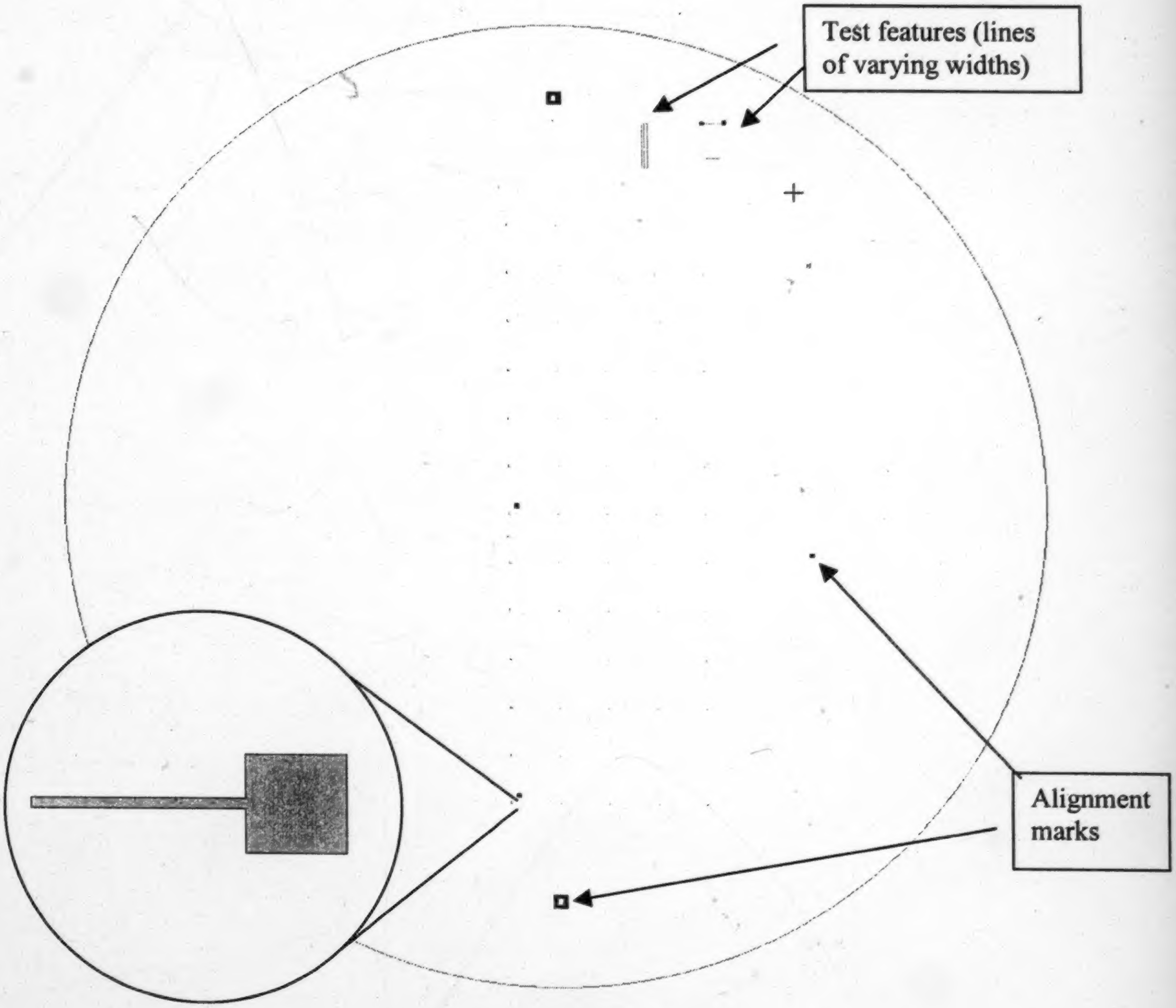

Test features (lines of varying widths)
Alignment marks

Figure A - 3 MIMSSF Mask

MIMSSF – Small Features Mask, written with a 4mm lens to minimal feature sizes of 1µm. The features themselves are essentially just the top electrode.

Appendix D: Harmonic Radar Tag Layout

A harmonic radar prototype tag was designed based on the MIM diode and an impedance matched folded dipole antenna. The dipole was designed for the 2GHz fundamental and 6GHz harmonic, and for the testing case of a 5cm SiO_2 base to simulate substrate coupling, which is expected to play a greater role as the device is scaled to smaller dimensions (higher frequencies).

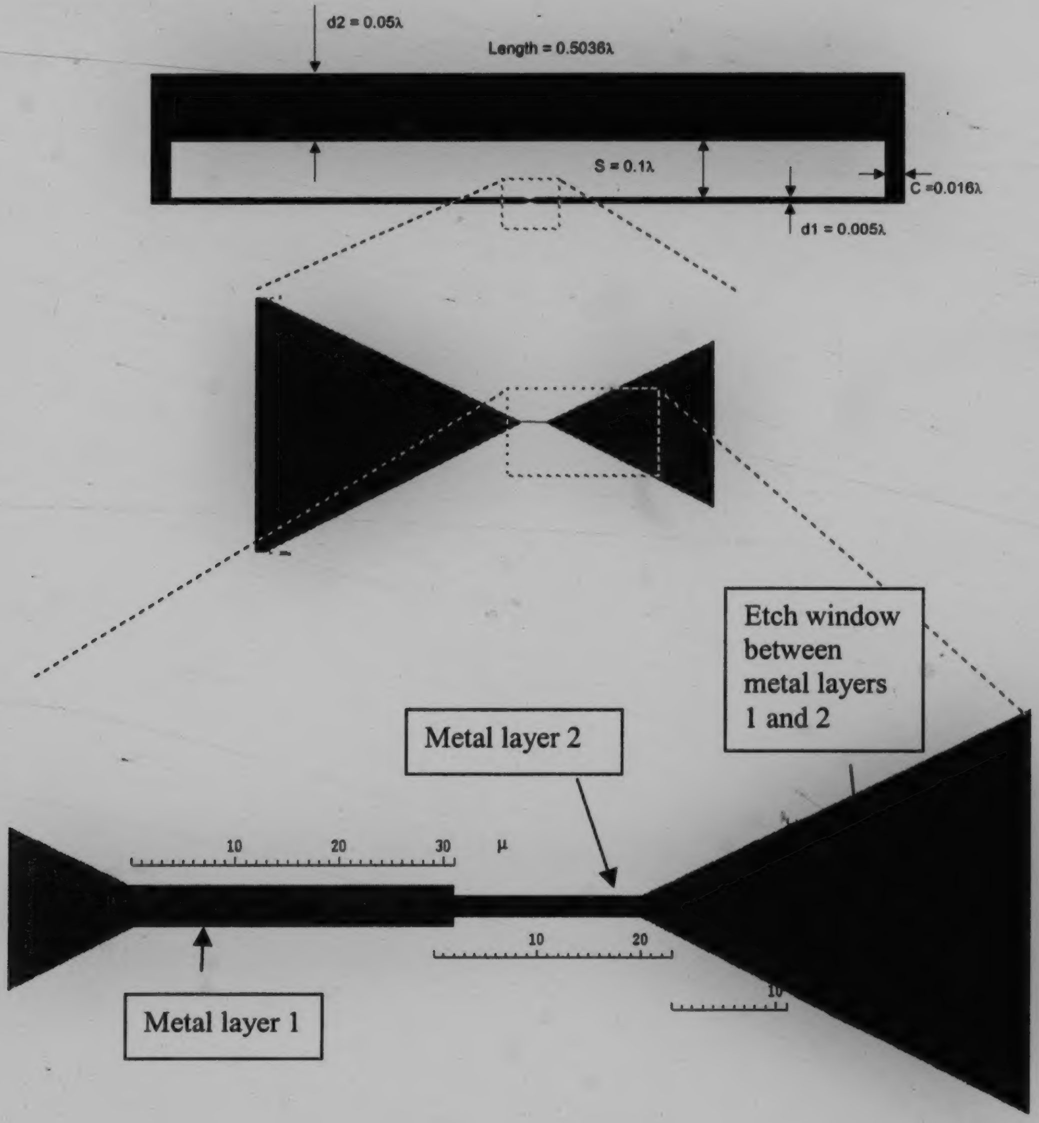

d2 = 0.05λ
Length = 0.5036λ
S = 0.1λ
C =0.016λ
d1 = 0.005λ
Etch window between metal layers 1 and 2
Metal layer 2
10
20
30
μ
10
20
10
Metal layer 1

Figure A - 4 Folded-Dipole Harmonic Radar Tag Antenna: The antenna is specified to the dimensions discussed in Chapter 8. The diode is a simple cross-layer configuration, similar to the MIM diode process.

Appendix E: Program Code

The source code for simulations used in this thesis is listed here. Electrostatic simulations using ANSYS Multiphysics for demonstration of MITT operation and curve fitting of measured MIM diode *i-v* data to theory in MatLab code are given.

MITT Electrostatic Analysis (ANSYS)

ANSYS/ED Multiphysics version 6.1 was used for the electrostatic simulation of the MITT tunnel junctions using the 2D electrostatic element PLANE121. Refer to page 124, in Chapter 7, for results of simulation.

```
! MITT.LGW
! by Jay Sulima, March, 2004
! Electrostatic Simulation of a Metal-Insulator Tunnelling Transistor
/PREP7
! Define Geometry
K,1,0,0,0,
K,2,200,0,0,
K,3,200,100,0,
K,4,0,100,0,
K,5,-50,0,0,
K,5,-50,-50,0,
K,6,250,-50,0,
K,7,250,0,0,
K,8,-50,0,0,
K,8,-50,0,0,
LSTR,8,1
LSTR,1,4
LSTR,4,3
LSTR,3,2
LSTR,2,1
LSTR,2,7
LSTR,7,6
LSTR,6,5
LSTR,5,8
FLST,2,4,4
FITEM,2,2
FITEM,2,3
FITEM,2,4
FITEM,2,5
AL,P51X
FLST,2,6,4
FITEM,2,1
FITEM,2,5
FITEM,2,6
FITEM,2,7
FITEM,2,8
FITEM,2,9
AL,P51X
FLST,2,2,5,ORDE,2
FITEM,2,1
```

```
FITEM,2,-2
! Glue areas together
AGLUE,P51X

! Define material properties
ET,1,PLANE121
KEYOPT,1,3,0
KEYOPT,1,4,0
KEYOPT,1,5,0
KEYOPT,1,7,0
MPTEMP,,,,,,,,
MPTEMP,1,0
MPDATA,PERX,1,,11.5
MPTEMP,,,,,,,,
MPTEMP,1,0
MPDATA,PERX,2,,4

CM,_Y,AREA
ASEL,,,,2
CM,_Y1,AREA
CMSEL,S,_Y
!*
CMSEL,S,_Y1
AATT,1,,1,0,
CMSEL,S,_Y
CMDELE,_Y
CMDELE,_Y1
!*
CM,_Y,AREA
ASEL, , , ,1
CM,_Y1,AREA
CMSEL,S,_Y
!*
CMSEL,S,_Y1
AATT,2, ,1,0,
CMSEL,S,_Y
CMDELE,_Y
CMDELE,_Y1

! Mesh areas
MSHKEY,0
FLST,5,2,5,ORDE,2
FITEM,5,1
FITEM,5,-2
CM,_Y,AREA
ASEL, , , ,P51X
CM,_Y1,AREA
CHKMSH,'AREA'
CMSEL,S,_Y
AMESH,_Y1
CMDELE,_Y
CMDELE,_Y1
CMDELE,_Y2
FLST,5,6,1,ORDE,6
FITEM,5,1
FITEM,5,18
FITEM,5,25
```

```
FITEM,5,37
FITEM,5,77
FITEM,5,-78
CM,_Y,NODE
NSEL, , , ,P51X
CM,_Y1,NODE
CMSEL,S,_Y
CMDELE,_Y
NREFINE,_Y1, , ,1,1,1,1
CMDELE,_Y1
FINISH

! Solve
/SOLU
ANTYPE,0
FLST,2,2,4,ORDE,2
FITEM,2,1
FITEM,2,-2
/GO
DL,P51X, ,VOLT,0
FLST,2,2,4,ORDE,2
FITEM,2,4
FITEM,2,6
/GO
DL,P51X, ,VOLT,5
FLST,2,1,4,ORDE,1
FITEM,2,8
/GO
DL,P51X, ,VOLT,2.5
SOLVE
FINISH
/POST1
! Show solution in graph
PLNSOL,VOLT, ,0,
```

Curve Fitting (MATLAB)

Two files are listed, Fitdata.m and FitJFN.m. FitJFN is used by Fitdata.m, which uses the MatLab function fmins() to minimize the error function returned by FitJFN. FitJFN sums the total difference (error) between measured data (extrapolated polynomial functions) and theoretical functions and returns the result to Fitdata.m. The FitJFN function is re-run with new input parameters (d, phi, JA encompassed in the lambda variable) by fmins() until a suitably close fit between data and theory is attained. The results of the code can be found on page 106, in Chapter 6.

```
% Fitdata.m
% by Jay Sulima September, 2003
% Using FitJFN.m and the fmins() minimization function, fits the i-v
% data to a quantum tunnelling model

global Data

V = [-2:0.11:2];
Vin = 1./V;
```

```
% Extrapolated data in polynomial function form
for i = 1:length(V)
   % 7M -2 to 2
   %y(i) = 13.881*V(i)*V(i)*V(i)-1.1442*V(i)*V(i)+9.5879*V(i)+0.1602;
   % 7L -2 to 2
   y(i) = 2.2599*V(i)*V(i)*V(i)+0.2612*V(i)*V(i)+12.761*V(i)-0.0463;
   % 7E -1 to 1
   %y(i) = 128.08*V(i)^5 - 37.298*V(i)^4 +109.58*V(i)^3 + 26.612*V(i)^2
+53.161*V(i) -1.4322;
   % 6I -5 to 5
   %y(i) = 0.0002*V(i)^6 + 0.0134*V(i)^5 -0.0029*V(i)^4 -0.0965*V(i)^3
+0.0389*V(i)^2 +2.6377*V(i) -0.0601;
   % 6D -7 to 7
   %y(i) = 0.0002*V(i)^6 - 0.0048*V(i)^5 -0.0114*V(i)^4 +0.5598*V(i)^3
+0.2528*V(i)^2 +13.255*V(i) -0.0829;
   % 4C -4 to 4
   %y(i) = -0.0048*V(i)^5 - 0.0007*V(i)^4 + 0.1754*V(i)^3 +
0.0535*V(i)^2 + 1.803*V(i) + 0.0203;
   % 6M from run 2 -0.75 to 0.75
   %y(i) = -0.6663*V(i)^6 - 5.4788*V(i)^5 + 0.7951*V(i)^4 +
11.013*V(i)^3 - 0.2206*V(i)^2 + 3.6848*V(i) - 0.0071;
   Data(i,1) = V(i);
end

% setup plotting
yx = log(y./(V.*V));
axis([0 max(Vin) 0 max(yx(i))])
global Plothandle
plot(V,y,'ro','EraseMode','xo')
xlabel('Voltage (V)');
ylabel('Current (μA)');
hold on;
Plothandle = plot(V,y,'EraseMode','xo');
axis([-max(V) max(V) -max(y(i)) max(y(i))])

% Have initial values changed? If not then use result values from last
% run for d, phi, JA
if ~ exist('lamold')
   lamold = [0 0 0 0];
end

if ~ exist('holdold')
   holdold = 1;
end

if exist('lam')
   holdold = lamold;
end

% start with d > 4Å, JA = 10^-11, phi = 0.1 to get best results
% lam/lambda is in the form [JA phi d pinhole-correction]
lam = [10^-11 0.1 4.5 0.1];
lamold = lam;
% use last values if the above values haven't been changed manually
since last run
if exist('lambda') & holdold == lamold
   lam = lambda;
```

```
end

trace = 0;
tol = 0.0000001;
lambda = fmins('fitJFN',lam,[trace tol]);

placenote = max(y(i))-0.1*max(y(i));
placenote2 = max(y(i))-0.2*max(y(i));
placenote3 = max(y(i))-0.3*max(y(i));
text(0.1,placenote,['JA  = ',num2str(abs(lambda(1)))]);
text(0.1,placenote2,['phi = ',num2str(lambda(2))]);
text(0.1,placenote3, ['d   = ',num2str(lambda(3))]);
lambda
hold off;

% ----------------------------------------------------------------------
% fitJFN.m
% by Jay Sulima, September, 2003
% The quantum tunnelling model used by FitData.m to find a
% theoretical fit for the given curves

function err = fitJFN(lambda)

global Data Plothandle

V = [-2:0.11:2];
Vin = 1./V;

% Extrapolated data in polynomial function form
for i = 1:length(V)
   % 7M -2 to 2
   %y(i) = 13.881*V(i)*V(i)*V(i)-1.1442*V(i)*V(i)+9.5879*V(i)+0.1602;
   % 7L -2 to 2
   y(i) = 2.2599*V(i)*V(i)*V(i)+0.2612*V(i)*V(i)+12.761*V(i)-0.0463;
   % 7E -1 to 1
   %y(i) = 128.08*V(i)^5 - 37.298*V(i)^4 +109.58*V(i)^3 + 26.612*V(i)^2
+53.161*V(i) -1.4322;
   % 6I -5 to 5
   %y(i) = 0.0002*V(i)^6 + 0.0134*V(i)^5 -0.0029*V(i)^4 -0.0965*V(i)^3
+0.0389*V(i)^2 +2.6377*V(i) -0.0601;
   % 6D -7 to 7
   %y(i) = 0.0002*V(i)^6 - 0.0048*V(i)^5 -0.0114*V(i)^4 +0.5598*V(i)^3
+0.2528*V(i)^2 +13.255*V(i) -0.0829;
   % 4C -4 to 4
   %y(i) = -0.0048*V(i)^5 - 0.0007*V(i)^4 + 0.1754*V(i)^3 +
0.0535*V(i)^2 + 1.803*V(i) + 0.0203;
   % 6M from run 2
   %y(i) = -0.6663*V(i)^6 - 5.4788*V(i)^5 + 0.7951*V(i)^4 +
11.013*V(i)^3 - 0.2206*V(i)^2 + 3.6848*V(i) - 0.0071;
end

JA          = lambda(1);      % Junction Area
phi         = lambda(2);      % barrier height
d           = lambda(3);      % barrier width
pinholes    = lambda(4);      % pinhole resistance, not usually needed
T = 300;                      % temperature in K
```

```
JA0 = 10^-8 * 10^6;           % conversion from cm² to µm² then for µA
J0fn        = 3.38 * 10^10;
A           = 1.025;          % constant used in Simmons model
B           = -0.689;         % constant used in Simmons model
KAlO        = 11.9;           % dielectric constant of AlO
KniO        = 4;              % dielectric constant of NiO
K           = KNiO;

% i-v without image forces
% See Ref. 9 for details

yf0 = (V./abs(V)).* abs(JA).* JA0.*3.16*10^10 .* abs(phi)^0.5 .*
(abs(V)./abs(d)).*exp(-1.025.*abs(d)*abs(phi)^0.5); %low voltage range
yfn = (V./abs(V)).*J0fn .* abs(JA).*JA0.* ((V.^2 )./(d*d*abs(phi))) .*
(exp(( B * abs(d) * abs(phi)^1.5)./abs(V)));%-
(1+2*V./abs(phi)).*exp(B*abs(phi).^1.5 *(1+2*V./abs(phi)).^0.5
*abs(d)./V));
%yf = (V./abs(V)).*6.2*10^10 .* 25.*JA JA0.*.* ( abs(phi)*exp(-A*abs(d)
* abs(phi)^0.5) - (abs(phi)+abs(V)).*exp(-
A*abs(d)*(abs(phi)+abs(V)).^0.5))/(d*d);
yf = (V./abs(V)).*6.2*10^10 .* abs(JA).*JA0.* ((abs(phi)-
abs(V)./2).*exp(-A*abs(d)*(abs(phi)-abs(V)./2).^0.5) -
(abs(phi)+abs(V)/2).*exp(-
A*abs(d)*(abs(phi)+abs(V)/2).^0.5))/(d*d);%intermediate

% i-v with image forces - most successful fits
% See Ref. 9 for details

s1lo = 6./(K*phi);
s2lo =  abs(d - s1lo);
s1mi = s1lo;
s2mi = d*(1-46./(3*phi*K*d+20-2*abs(V)*K*d))+6./(K*phi);
s1hi = s1lo;
s2hi = (phi*K*d-28)./(K*abs(V));

phiIlo = phi-(5.75./(K*(s2lo-s1lo)))*log(s2lo*(d-s1lo)./(s1lo*(d-
s2lo)));
phiImi = phi-(abs(V)./(2*d)).*(s1mi+s2mi)-(5.75./(K.*(s2mi-
s1mi))).*log(s2mi.*abs(d-s1mi)./(s1mi.*abs(d-s2mi)));
phiIhi = phi-(abs(V)./(2*d)).*(s1hi+s2hi)-(5.75./(K.*(s2hi-
s1hi))).*log(s2hi.*abs(d-s1hi)./(s1hi.*abs(d-s2hi)));

delslo = s2lo - s1lo;
delsmi = s2mi - s1mi;
delshi = s2hi - s1hi;

ylo = JA*((3.16*10^10)./delslo).*(phiIlo).^0.5 .* V .* exp(-
1.025*delslo.*phiIlo.^0.5);
ymi = JA*(V./abs(V)).*(6.2*10^10./(delsmi).^2).*(abs(phiImi).*exp(-
A.*delsmi.*abs(phiImi).^0.5)-(abs(phiImi)+abs(V)).*exp(-
A.*delsmi.*(abs(phiImi)+abs(V)).^0.5));
yhi = JA*(V./abs(V)).*(6.2*10^10./(delshi).^2).*(abs(phiIhi).*exp(-
A.*delshi.*abs(phiIhi).^0.5)-(abs(phiIhi)+abs(V)).*exp(-
A.*delshi.*(abs(phiIhi)+abs(V)).^0.5));
err = 0;

% Sanchez model
% See Ref. 112 for details

phi = abs(phi);
JA  = abs(JA);
areaum = JA0.*JA;
S = 1.025*phi*d^0.5;
m = S/(24*phi);
n = (1/6)*(S/(4*phi))^2;
Rd = S*exp(S)/(324*phi*areaum);
yfSan = (V+m*V.^2+n*V.^3)./Rd;

% Thermionic current and thermal distribution adjustments
% thermionic emission @ 300K

phin  = phi - V/2 - 10/d;
yth = JA*JA0*9*10^6*10.^(-16.82*phin);

Jtherm = (V./abs(V)).*JA.*9.*10.^(-16.82*(abs(phi)-abs(V)/2.-10/d)); %
additive term
%Jtherm = 1+3.*10^-9.*d.*d.*T^2./(phi-V/2);
% multiplicative term
JthermFN  = 1+ 6*10^-9.*d*d*T^2.*phi./V;

% Adjust the following to fit with each model
for i=1:length(V)
  if abs(V(i)) > abs(phi)           % high voltage region
     yf(i) = yhi(i) * JthermFN(i);
  elseif abs(V(i)) < abs(phi)/50;   % small voltage region
    yf(i) = ylo(i)  ;
     yf(i) = yf(i) * JthermFN(i);
  else                              % mid-voltage region
     yf(i) = ymi(i) * JthermFN(i);
  end
  yf(i) = yf(i) + V(i)./pinholes;
end

set(Plothandle,'ydata',yf);
drawnow;

%error function
err = 0;
for i = 1:length(V)
   err = err + abs(yf(i)*yf(i) - y(i)*y(i))/(abs(y(i)));
   err = norm(yf-y);
end
```